

# Scaling, Proximity, and Optimization of Integrally Convex Functions\*

Satoko Moriguchi,<sup>†</sup> Kazuo Murota,<sup>‡</sup> Akihisa Tamura,<sup>§</sup> Fabio Tardella<sup>¶</sup>

March 29, 2017

## Abstract

In discrete convex analysis, the scaling and proximity properties for the class of  $L^1$ -convex functions were established more than a decade ago and have been used to design efficient minimization algorithms. For the larger class of integrally convex functions of  $n$  variables, we show here that the scaling property only holds when  $n \leq 2$ , while a proximity theorem can be established for any  $n$ , but only with a superexponential bound. This is, however, sufficient to extend the classical logarithmic complexity result for minimizing a discrete convex function of one variable to the case of integrally convex functions of any fixed number of variables.

## 1 Introduction

The proximity-scaling approach is a fundamental technique in designing efficient algorithms for discrete or combinatorial optimization. For a function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  in integer variables and a positive integer  $\alpha$ , called scaling unit, the  $\alpha$ -scaling of  $f$  means the function  $f^\alpha$  defined by  $f^\alpha(x) = f(\alpha x)$  ( $x \in \mathbb{Z}^n$ ). A proximity theorem is a result guaranteeing that a (local) minimum of the scaled function  $f^\alpha$  is close to a minimizer of the original function  $f$ . The scaled function  $f^\alpha$  is simpler in shape, and hence easier to minimize, whereas the quality of the obtained minimizer of  $f^\alpha$  as an approximation to the minimizer of  $f$  is guaranteed by a proximity theorem. The proximity-scaling approach consists in applying this idea for a decreasing sequence of  $\alpha$ , often by halving the scale unit  $\alpha$ . A generic form of a proximity-scaling algorithm may be described as follows, where  $K_\infty$  denotes the  $\ell_\infty$ -size of the effective domain of  $f$ , and  $B(n, \alpha)$  denotes the proximity bound in  $\ell_\infty$ -distance.

### Proximity-scaling algorithm

S0: Find an initial vector  $x$  with  $f(x) < +\infty$ , and set  $\alpha := 2^{\lceil \log_2 K_\infty \rceil}$ .

S1: Find a vector  $y$  with  $\|\alpha y\|_\infty \leq B(n, \alpha)$  that is a (local) minimizer of  $\tilde{f}(y) = f(x + \alpha y)$ ,

---

\*The extended abstract of this paper is included in the Proceedings of the 27th International Symposium on Algorithms and Computation (ISAAC), Sydney, December 12–14, 2016. Leibniz International Proceedings in Informatics (LIPIcs), **64** (2016), 57:1–57:12, Dagstuhl Publishing.

<sup>†</sup>Tokyo Metropolitan University, satoko5@tmu.ac.jp

<sup>‡</sup>Tokyo Metropolitan University, murota@tmu.ac.jp

<sup>§</sup>Keio University, aki-tamura@math.keio.ac.jp

<sup>¶</sup>Sapienza University of Rome, fabio.tardella@uniroma1.it

and set  $x := x + \alpha y$ .

S2: If  $\alpha = 1$ , then stop ( $x$  is a minimizer of  $f$ ).

S3: Set  $\alpha := \alpha/2$ , and go to S1.

The algorithm consists of  $O(\log_2 K_\infty)$  scaling phases. This approach has been particularly successful for resource allocation problems [8, 9, 10, 16] and for convex network flow problems (under the name of “capacity scaling”) [1, 14, 15]. Different types of proximity theorems have also been investigated: proximity between integral and real optimal solutions [9, 31, 32], among others. For other types of algorithms of nonlinear integer optimization, see, e.g., [5].

In discrete convex analysis [22, 23, 24, 25], a variety of discrete convex functions are considered. A *separable convex* function is a function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  that can be represented as  $f(x) = \varphi_1(x_1) + \dots + \varphi_n(x_n)$ , where  $x = (x_1, \dots, x_n)$ , with univariate discrete convex functions  $\varphi_i : \mathbb{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  satisfying  $\varphi_i(t-1) + \varphi_i(t+1) \geq 2\varphi_i(t)$  for all  $t \in \mathbb{Z}$ .

A function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is called *integrally convex* if its local convex extension  $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is (globally) convex in the ordinary sense, where  $\tilde{f}$  is defined as the collection of convex extensions of  $f$  in each unit hypercube  $\{x \in \mathbb{R}^n \mid a_i \leq x_i \leq a_i + 1 \ (i = 1, \dots, n)\}$  with  $a \in \mathbb{Z}^n$ ; see Section 2 for precise statements.

A function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is called  *$L^h$ -convex* if it satisfies one of the equivalent conditions in Theorem 1.1 below. For  $x, y \in \mathbb{Z}^n$ ,  $x \vee y$  and  $x \wedge y$  denote the vectors of componentwise maximum and minimum of  $x$  and  $y$ , respectively. Discrete midpoint convexity of  $f$  for  $x, y \in \mathbb{Z}^n$  means

$$f(x) + f(y) \geq f\left(\left\lceil \frac{x+y}{2} \right\rceil\right) + f\left(\left\lfloor \frac{x+y}{2} \right\rfloor\right), \quad (1.1)$$

where  $\lceil \cdot \rceil$  and  $\lfloor \cdot \rfloor$  denote the integer vectors obtained by componentwise rounding-up and rounding-down to the nearest integers, respectively. The effective domain of a function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is the set  $\text{dom } f = \{x \in \mathbb{Z}^n \mid f(x) < +\infty\}$ . We use the notation  $\mathbf{1} = (1, 1, \dots, 1)$  and  $\mathbf{1}_i$  for the  $i$ -th unit vector  $(0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$ , with the convention  $\mathbf{1}_0 = \mathbf{0}$ .

**Theorem 1.1** ([2, 4, 23]). *For  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  the following conditions, (a) to (d), are equivalent:<sup>1</sup>*

(a)  *$f$  is integrally convex and submodular:*

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \quad (x, y \in \mathbb{Z}^n). \quad (1.2)$$

(b)  *$f$  satisfies discrete midpoint convexity (1.1) for all  $x, y \in \mathbb{Z}^n$ .*

(c)  *$f$  satisfies discrete midpoint convexity (1.1) for all  $x, y \in \mathbb{Z}^n$  with  $\|x - y\|_\infty \leq 2$ , and the effective domain has the property:  $x, y \in \text{dom } f \Rightarrow \lceil (x+y)/2 \rceil, \lfloor (x+y)/2 \rfloor \in \text{dom } f$ .*

(d)  *$f$  satisfies translation-submodularity:*

$$f(x) + f(y) \geq f((x - \mu \mathbf{1}) \vee y) + f(x \wedge (y + \mu \mathbf{1})) \quad (x, y \in \mathbb{Z}^n, \ 0 \leq \mu \in \mathbb{Z}). \quad (1.3)$$

■

A function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  is called  *$M^h$ -convex* if it satisfies an exchange property: For any  $x, y \in \text{dom } f$  and any  $i \in \text{supp}^+(x - y)$ , there exists  $j \in \text{supp}^-(x - y) \cup \{0\}$  such that

$$f(x) + f(y) \geq f(x - \mathbf{1}_i + \mathbf{1}_j) + f(y + \mathbf{1}_i - \mathbf{1}_j), \quad (1.4)$$

---

<sup>1</sup> $\mathbb{Z}$ -valued functions are treated in [4, Theorem 3], but the proof is valid for  $\mathbb{R}$ -valued functions.

where, for  $z \in \mathbb{Z}^n$ ,  $\text{supp}^+(z) = \{i \mid z_i > 0\}$  and  $\text{supp}^-(z) = \{j \mid z_j < 0\}$ . It is known (and easy to see) that a function is separable convex if and only if it is both  $L^h$ -convex and  $M^h$ -convex.

Integrally convex functions constitute a common framework for discrete convex functions, including separable convex,  $L^h$ -convex and  $M^h$ -convex functions as well as  $L_2^h$ -convex and  $M_2^h$ -convex functions [23], and BS-convex and UJ-convex functions [3]. The concept of integral convexity is used in formulating discrete fixed point theorems [11, 12, 35], and designing solution algorithms for discrete systems of nonlinear equations [17, 34]. In game theory the integral concavity of payoff functions guarantees the existence of a pure strategy equilibrium in finite symmetric games [13].

The scaling operation preserves  $L^h$ -convexity, that is, if  $f$  is  $L^h$ -convex, then  $f^\alpha$  is  $L^h$ -convex.  $M^h$ -convexity is subtle in this respect: for an  $M^h$ -convex function  $f$ ,  $f^\alpha$  remains  $M^h$ -convex if  $n \leq 2$ , while this is not always the case if  $n \geq 3$ .

**Example 1.1.** Here is an example to show that  $M^h$ -convexity is not preserved under scaling. Let  $f$  be the indicator function of the set  $S = \{c_1(1, 0, -1) + c_2(1, 0, 0) + c_3(0, 1, -1) + c_4(0, 1, 0) \mid c_i \in \{0, 1\}\} \subseteq \mathbb{Z}^3$ . Then  $f$  is an  $M^h$ -convex function, but  $f^2 (=f^\alpha$  with  $\alpha = 2$ ), being the indicator function of  $\{(0, 0, 0), (1, 1, -1)\}$ , is not  $M^h$ -convex. (In other words,  $S$  is an  $M^h$ -convex set, and  $\{(0, 0, 0), (1, 1, -1)\}$  is not.) This example is a reformulation of [23, Note 6.18] for  $M$ -convex functions to  $M^h$ -convex functions. ■

It is rather surprising that nothing is known about scaling for integrally convex functions. Example 1.1 does not demonstrate the lack of scaling property of integrally convex functions, since  $f^2$  above is integrally convex, though not  $M^h$ -convex.

As for proximity theorems, the following facts are known for separable convex,  $L^h$ -convex and  $M^h$ -convex functions. In the following three theorems we assume that  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $\alpha$  is a positive integer, and  $x^\alpha \in \text{dom } f$ . It is noteworthy that the proximity bound is independent of  $n$  for separable convex functions, and coincides with  $n(\alpha - 1)$ , which is linear in  $n$ , for  $L^h$ -convex and  $M^h$ -convex functions.

**Theorem 1.2.** *Suppose that  $f$  is a separable convex function. If  $f(x^\alpha) \leq f(x^\alpha + \alpha d)$  for all  $d \in \{\mathbf{1}_i, -\mathbf{1}_i \mid (1 \leq i \leq n)\}$ , then there exists a minimizer  $x^*$  of  $f$  with  $\|x^\alpha - x^*\|_\infty \leq \alpha - 1$ . ■*

**Theorem 1.3** ([15, 23]). *Suppose that  $f$  is an  $L^h$ -convex function. If  $f(x^\alpha) \leq f(x^\alpha + \alpha d)$  for all  $d \in \{0, 1\}^n \cup \{0, -1\}^n$ , then there exists a minimizer  $x^*$  of  $f$  with  $\|x^\alpha - x^*\|_\infty \leq n(\alpha - 1)$ . ■*

**Theorem 1.4** ([18, 23]). *Suppose that  $f$  is an  $M^h$ -convex function. If  $f(x^\alpha) \leq f(x^\alpha + \alpha d)$  for all  $d \in \{\mathbf{1}_i, -\mathbf{1}_i \mid (1 \leq i \leq n), \mathbf{1}_i - \mathbf{1}_j \mid (i \neq j)\}$ , then there exists a minimizer  $x^*$  of  $f$  with  $\|x^\alpha - x^*\|_\infty \leq n(\alpha - 1)$ . ■*

Based on the above results and their variants, efficient algorithms for minimizing  $L^h$ -convex and  $M^h$ -convex functions have been successfully designed with the proximity-scaling approach ([18, 20, 21, 23, 30, 33]). Proximity theorems are also available for  $L_2^h$ -convex and  $M_2^h$ -convex functions [27] and  $L$ -convex functions on graphs [6, 7]. However, no proximity theorem has yet been proved for integrally convex functions.

The new findings of this paper are

- A “box-barrier property” (Theorem 2.4), which allows us to restrict the search for a global minimum;

- Stability of integral convexity under scaling when  $n = 2$  (Theorem 3.3), and an example to demonstrate its failure when  $n \geq 3$  (Example 3.1);
- A proximity theorem with a superexponential bound  $[(n + 1)!/2^{n-1}](\alpha - 1)$  for all  $n$  (Theorem 5.1), and the impossibility of finding a proximity bound of the form  $B(n)(\alpha - 1)$  where  $B(n)$  is linear or smaller than quadratic (Examples 4.4 and 4.5).

As a consequence of our proximity and scaling results, we derive that:

- When  $n$  is fixed, an integrally convex function can be minimized in  $O(\log_2 K_\infty)$  time by standard proximity-scaling algorithms, where  $K_\infty = \max\{\|x - y\|_\infty \mid x, y \in \text{dom } f\}$  denotes the  $\ell_\infty$ -size of  $\text{dom } f$ .

This paper is organized as follows. In Section 2 the concept of integrally convex functions is reviewed with some new observations. Their scaling property is clarified in Section 3. After a preliminary discussion in Section 4, a proximity theorem for integrally convex functions is established in Section 5. Algorithmic implications of the proximity-scaling results are discussed in Section 6 and concluding remarks are made in Section 7.

## 2 Integrally Convex Functions

For  $x \in \mathbb{R}^n$  the integer neighborhood of  $x$  is defined as

$$N(x) = \{z \in \mathbb{Z}^n \mid |x_i - z_i| < 1 \ (i = 1, \dots, n)\}.$$

For a function  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  the local convex extension  $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  of  $f$  is defined as the union of all convex envelopes of  $f$  on  $N(x)$  as follows:

$$\tilde{f}(x) = \min \left\{ \sum_{y \in N(x)} \lambda_y f(y) \mid \sum_{y \in N(x)} \lambda_y y = x, (\lambda_y) \in \Lambda(x) \right\} \quad (x \in \mathbb{R}^n), \quad (2.1)$$

where  $\Lambda(x)$  denotes the set of coefficients for convex combinations indexed by  $N(x)$ :

$$\Lambda(x) = \{(\lambda_y \mid y \in N(x)) \mid \sum_{y \in N(x)} \lambda_y = 1, \lambda_y \geq 0 \ (\forall y \in N(x))\}.$$

If  $\tilde{f}$  is convex on  $\mathbb{R}^n$ , then  $f$  is said to be *integrally convex*. A set  $S \subseteq \mathbb{Z}^n$  is said to be *integrally convex* if the convex hull  $\overline{S}$  of  $S$  coincides with the union of the convex hulls of  $S \cap N(x)$  over  $x \in \mathbb{R}^n$ , i.e., if, for any  $x \in \mathbb{R}^n$ ,  $x \in \overline{S}$  implies  $x \in \overline{S \cap N(x)}$ .

Integral convexity is preserved under the operations of origin shift, permutation of components, and componentwise (individual) sign inversion. For later reference we state these facts as a proposition.

**Proposition 2.1.** *Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function.*

- (1) *For any  $z \in \mathbb{Z}^n$ ,  $f(z + x)$  is integrally convex in  $x$ .*
- (2) *For any permutation  $\sigma$  of  $(1, 2, \dots, n)$ ,  $f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$  is integrally convex in  $x$ .*
- (3) *For any  $s_1, s_2, \dots, s_n \in \{+1, -1\}$ ,  $f(s_1 x_1, s_2 x_2, \dots, s_n x_n)$  is integrally convex in  $x$ . ■*

Integral convexity of a function can be characterized by a local condition under the assumption that the effective domain is an integrally convex set. The following theorem is proved in [2] when the effective domain is an integer interval (discrete rectangle). An alternative proof, which is also valid for the general case, is given in Appendix A.

**Theorem 2.2** ([2, Proposition 3.3]). *Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be a function with an integrally convex effective domain. Then the following properties are equivalent:*

- (a)  *$f$  is integrally convex.*
- (b) *For every  $x, y \in \text{dom } f$  with  $\|x - y\|_\infty = 2$  we have*

$$\tilde{f}\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f(x) + f(y)). \quad (2.2)$$

■

**Theorem 2.3** ([2, Proposition 3.1]; see also [23, Theorem 3.21]). *Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function and  $x^* \in \text{dom } f$ . Then  $x^*$  is a minimizer of  $f$  if and only if  $f(x^*) \leq f(x^* + d)$  for all  $d \in \{-1, 0, +1\}^n$ .*

■

The local characterization of global minima stated in Theorem 2.3 above can be generalized to the following form.

**Theorem 2.4** (Box-barrier property). *Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function, and let  $p \in (\mathbb{Z} \cup \{-\infty\})^n$  and  $q \in (\mathbb{Z} \cup \{+\infty\})^n$ , where  $p \leq q$ . Define*

$$\begin{aligned} S &= \{x \in \mathbb{Z}^n \mid p_i < x < q_i \ (i = 1, \dots, n)\}, \\ W_i^+ &= \{x \in \mathbb{Z}^n \mid x_i = q_i, \ p_j \leq x_j \leq q_j \ (j \neq i)\} \quad (i = 1, \dots, n), \\ W_i^- &= \{x \in \mathbb{Z}^n \mid x_i = p_i, \ p_j \leq x_j \leq q_j \ (j \neq i)\} \quad (i = 1, \dots, n), \end{aligned}$$

and  $W = \bigcup_{i=1}^n (W_i^+ \cup W_i^-)$ . *Let  $\hat{x} \in S \cap \text{dom } f$ . If  $f(\hat{x}) \leq f(y)$  for all  $y \in W$ , then  $f(\hat{x}) \leq f(z)$  for all  $z \in \mathbb{Z}^n \setminus S$ .*

*Proof.* Let  $U = \bigcup_{i=1}^n \{x \in \mathbb{R}^n \mid x_i \in \{p_i, q_i\}, \ p_j \leq x_j \leq q_j \ (j \neq i)\}$ , for which we have  $U \cap \mathbb{Z}^n = W$ . For a point  $z \in \mathbb{Z}^n \setminus S$ , the line segment connecting  $\hat{x}$  and  $z$  intersects  $U$  at a point, say,  $u \in \mathbb{R}^n$ . Then its integral neighborhood  $N(u)$  is contained in  $W$ . Since the local convex extension  $\tilde{f}(u)$  is a convex combination of the  $f(y)$ 's with  $y \in N(u)$ , and  $f(y) \geq f(\hat{x})$  for every  $y \in W$ , we have  $\tilde{f}(u) \geq f(\hat{x})$ . On the other hand, it follows from integral convexity that  $\tilde{f}(u) \leq (1-\lambda)f(\hat{x}) + \lambda f(z)$  for some  $\lambda$  with  $0 < \lambda \leq 1$ . Hence  $f(\hat{x}) \leq \tilde{f}(u) \leq (1-\lambda)f(\hat{x}) + \lambda f(z)$ , and therefore,  $f(\hat{x}) \leq f(z)$ . □

Theorem 2.3 is a special case of Theorem 2.4 with  $p = \hat{x} - \mathbf{1}$  and  $q = \hat{x} + \mathbf{1}$ . Another special case of Theorem 2.4 with  $p_j = -\infty$  ( $j = 1, \dots, n$ ) and  $q_j = +\infty$  ( $j \neq i$ ) for a particular  $i$  takes the following form, which we use in Section 5.3.

**Corollary 2.5** (Hyperplane-barrier property). *Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function. Let  $\hat{x} \in \text{dom } f$ ,  $q \in \mathbb{Z}$ , and let  $i$  be an integer with  $1 \leq i \leq n$ . If  $\hat{x}_i < q$  and  $f(\hat{x}) \leq f(y)$  for all  $y \in \mathbb{Z}^n$  with  $y_i = q$ , then  $f(\hat{x}) \leq f(z)$  for all  $z \in \mathbb{Z}^n$  with  $z_i \geq q$ .*

■

We denote the sets of nonnegative integers and positive integers by  $\mathbb{Z}_+$  and  $\mathbb{Z}_{++}$ , respectively. For vectors  $a, b \in \mathbb{R}^n$  with  $a \leq b$ ,  $[a, b]_{\mathbb{R}}$  denotes the interval between  $a$  and  $b$ , i.e.,  $[a, b]_{\mathbb{R}} = \{x \in \mathbb{R}^n \mid a \leq x \leq b\}$ , and  $[a, b]_{\mathbb{Z}}$  the integer interval between  $a$  and  $b$ , i.e.,  $[a, b]_{\mathbb{Z}} = \{x \in \mathbb{Z}^n \mid a \leq x \leq b\}$ .

### 3 The Scaling Operation for Integrally Convex Functions

In this section we consider the scaling operation for integrally convex functions. Recall that, for  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  and  $\alpha \in \mathbb{Z}_{++}$ , the  $\alpha$ -scaling of  $f$  is defined to be a function  $f^\alpha : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  given by  $f^\alpha(x) = f(\alpha x)$  ( $x \in \mathbb{Z}^n$ ).

When  $n = 2$ , integral convexity is preserved under scaling. We first deal with integrally convex sets.

**Proposition 3.1.** *A set  $S \subseteq \mathbb{Z}^2$  is an integrally convex set if and only if it can be represented as  $S = \{(x_1, x_2) \in \mathbb{Z}^2 \mid a_i x_1 + b_i x_2 \leq c_i \ (i = 1, \dots, m)\}$  for some  $a_i, b_i \in \{-1, 0, 1\}$  and  $c_i \in \mathbb{Z}$  ( $i = 1, \dots, m$ ).* ■

We note that in the special case where all inequalities  $a_i x_1 + b_i x_2 \leq c_i$  ( $i = 1, \dots, m$ ) defining  $S$  in Proposition 3.1 satisfy the additional property  $a_i b_i \leq 0$ , the set  $S$  is actually an  $L^\natural$ -convex set [23], which is a special type of sublattice [28].

**Proposition 3.2.** *Let  $S \subseteq \mathbb{Z}^2$  be an integrally convex set and  $\alpha \in \mathbb{Z}_{++}$ . Then  $S^\alpha = \{x \in \mathbb{Z}^2 \mid \alpha x \in S\}$  is an integrally convex set.*

*Proof.* This follows easily from  $S^\alpha = (S \cap (\alpha\mathbb{Z})^2)/\alpha$  and Proposition 3.1. □

Next we turn to integrally convex functions.

**Theorem 3.3.** *Let  $f : \mathbb{Z}^2 \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function and  $\alpha \in \mathbb{Z}_{++}$ . Then the scaled function  $f^\alpha$  is integrally convex.*

*Proof.* The effective domain  $\text{dom } f^\alpha = (\text{dom } f \cap (\alpha\mathbb{Z})^2)/\alpha$  is an integrally convex set by Proposition 3.2. By Theorem 2.2 we only have to check condition (2.2) for  $f^\alpha$  with  $x = (0, 0)$  and  $y = (2, 0), (2, 2), (2, 1)$  (Proposition 2.1). That is, we are to show

$$f(0, 0) + f(2\alpha, 0) \geq 2f(\alpha, 0), \quad (3.1)$$

$$f(0, 0) + f(2\alpha, 2\alpha) \geq 2f(\alpha, \alpha), \quad (3.2)$$

$$f(0, 0) + f(2\alpha, \alpha) \geq f(\alpha, \alpha) + f(\alpha, 0). \quad (3.3)$$

The first two inequalities, (3.1) and (3.2), follow easily from integral convexity of  $f$ , whereas (3.3) is a special case of the basic parallelogram inequality (3.4) below with  $a = b = \alpha$ . □

**Proposition 3.4** (Basic parallelogram inequality). *For an integrally convex function  $f : \mathbb{Z}^2 \rightarrow \mathbb{R} \cup \{+\infty\}$  we have*

$$f(0, 0) + f(a + b, a) \geq f(a, a) + f(b, 0) \quad (a, b \in \mathbb{Z}_+). \quad (3.4)$$

*Proof.* We may assume  $a, b \geq 1$  and  $\{(0, 0), (a+b, a)\} \subseteq \text{dom } f$ , since otherwise the inequality (3.4) is trivially true. Then  $k(1, 1) + l(1, 0) \in \text{dom } f$  for all  $(k, l)$  with  $0 \leq k \leq a$  and  $0 \leq l \leq b$  by Proposition 3.1. We use the notation  $f_x(z) = f(x + z)$ . For each  $x \in \text{dom } f$  we have

$$f_x(0, 0) + f_x(2, 1) \geq f_x(1, 1) + f_x(1, 0)$$

by integral convexity of  $f$ . By adding these inequalities for  $x = k(1, 1) + l(1, 0)$  with  $0 \leq k \leq a - 1$  and  $0 \leq l \leq b - 1$ , we obtain (3.4). Note that all the terms involved in these inequalities are finite, since  $k(1, 1) + l(1, 0) \in \text{dom } f$  for all  $k$  and  $l$ . □



If  $n \geq 3$ ,  $f^\alpha$  is not always integrally convex. This is demonstrated by the following example.

**Example 3.1.** Consider the integrally convex function  $f : \mathbb{Z}^3 \rightarrow \mathbb{R} \cup \{+\infty\}$  defined on  $\text{dom } f = [(0, 0, 0), (4, 2, 2)]_{\mathbb{Z}}$  by

$x_2$	$f(x_1, x_2, 0)$					$x_2$	$f(x_1, x_2, 1)$					$x_2$	$f(x_1, x_2, 2)$				
2	3	1	1	1	3	2	2	1	0	0	0	2	3	2	1	0	0
1	1	0	0	0	0	1	1	0	0	0	0	1	2	1	0	0	0
0	0	0	0	0	3	0	0	0	0	0	0	0	3	0	0	0	3
	0	1	2	3	4		0	1	2	3	4		0	1	2	3	4
	$x_1$						$x_1$						$x_1$				

For the scaling with  $\alpha = 2$ , we have a failure of integral convexity:  $f(0, 0, 0) + f(4, 2, 2) < \min[f(2, 2, 2) + f(2, 0, 0), f(2, 2, 0) + f(2, 0, 2)]$ . The set  $S = \arg \min f = \{x \mid f(x) = 0\}$  is an integrally convex set, and  $S^\alpha = \{x \mid \alpha x \in S\} = \{(0, 0, 0), (1, 0, 0), (1, 0, 1), (2, 1, 1)\}$  is not an integrally convex set. ■

In view of the fact that the class of  $L^{\natural}$ -convex functions is stable under scaling, while this is not true for the superclass of integrally convex functions, we are naturally led to the question of finding an intermediate class of functions that is stable under scaling. See Section 7 for this issue.

## 4 Preliminary Discussion on Proximity Theorems

Before presenting a proximity theorem for integrally convex functions in Section 5, we establish in this section lower bounds for the proximity distance. We also present a proximity theorem for  $n = 2$ , as the proof is fairly simple in this particular case, though the proof method does not extend to general  $n \geq 3$ .

### 4.1 Lower bounds for the proximity distance

The following examples provide us with lower bounds for the proximity distance. The first three demonstrate the tightness of the bounds for separable convex functions,  $L^{\natural}$ -convex and  $M^{\natural}$ -convex functions given in Theorems 1.2, 1.3 and 1.4, respectively.

**Example 4.1** (Separable convex function). Let  $\varphi(t) = \max(-t, (\alpha - 1)(t - \alpha))$  for  $t \in \mathbb{Z}$  and define  $f(x) = \varphi(x_1) + \dots + \varphi(x_n)$ , which is separable convex. This function has a unique minimizer at  $x^* = (\alpha - 1, \dots, \alpha - 1)$ , whereas  $x^\alpha = \mathbf{0}$  is  $\alpha$ -local minimal and  $\|x^\alpha - x^*\|_\infty = \alpha - 1$ . This shows the tightness of the bound  $\alpha - 1$  given in Theorem 1.2. ■

**Example 4.2** ( $L^{\natural}$ -convex function). Consider  $X \subseteq \mathbb{Z}^n$  defined by

$$\begin{aligned} X &= \{x \in \mathbb{Z}^n \mid 0 \leq x_i - x_{i+1} \leq \alpha - 1 \ (i = 1, \dots, n-1), 0 \leq x_n \leq \alpha - 1\} \\ &= \{x \in \mathbb{Z}^n \mid x = \sum_{i=1}^n \mu_i \mathbf{1}_{\{1, 2, \dots, i\}}, \ 0 \leq \mu_i \leq \alpha - 1 \ (i = 1, \dots, n)\}, \end{aligned}$$

where  $\mathbf{1}_{\{1, 2, \dots, i\}} = (\overbrace{1, 1, \dots, 1}^i, 0, 0, \dots, 0)$ . The function  $f$  defined by  $f(x) = -x_1$  on  $\text{dom } f = X$  is an  $L^{\natural}$ -convex function and has a unique minimizer at  $x^* = (n(\alpha - 1), (n-1)(\alpha - 1), \dots, 2(\alpha - 1))$ .

1),  $\alpha - 1$ ). On the other hand,  $x^\alpha = \mathbf{0}$  is  $\alpha$ -local minimal, since  $X \cap \{-\alpha, 0, \alpha\}^n = \{\mathbf{0}\}$ . We have  $\|x^\alpha - x^*\|_\infty = n(\alpha - 1)$ , which shows the tightness of the bound  $n(\alpha - 1)$  given in Theorem 1.3. This example is a reformulation of [26, Remark 2.3] for L-convex functions to L<sup>b</sup>-convex functions. ■

**Example 4.3** (M<sup>b</sup>-convex function). Consider  $X \subseteq \mathbb{Z}^n$  defined by

$$\begin{aligned} X &= \{x \in \mathbb{Z}^n \mid 0 \leq x_1 + x_2 + \cdots + x_n \leq \alpha - 1, -(\alpha - 1) \leq x_i \leq 0 \ (i = 2, \dots, n)\} \\ &= \{x \in \mathbb{Z}^n \mid x = (\mu_1 + \mu_2 + \cdots + \mu_n, -\mu_2, -\mu_3, \dots, -\mu_n), \ 0 \leq \mu_i \leq \alpha - 1 \ (i = 1, \dots, n)\}. \end{aligned}$$

The function  $f$  defined by  $f(x) = -x_1$  on  $\text{dom } f = X$  is an M<sup>b</sup>-convex function and has a unique minimizer at  $x^* = (n(\alpha - 1), -(\alpha - 1), -(\alpha - 1), \dots, -(\alpha - 1))$ . On the other hand,  $x^\alpha = \mathbf{0}$  is  $\alpha$ -local minimal, since  $X \cap \{-\alpha, 0, \alpha\}^n = \{\mathbf{0}\}$ . We have  $\|x^\alpha - x^*\|_\infty = n(\alpha - 1)$ , which shows the tightness of the bound  $n(\alpha - 1)$  given in Theorem 1.4. This example is a reformulation of [26, Remark 2.8] for M-convex functions to M<sup>b</sup>-convex functions. ■

For integrally convex functions with  $n \geq 3$ , the bound  $n(\alpha - 1)$  is no longer valid. This is demonstrated by the following examples.

**Example 4.4.** Consider an integrally convex function  $f : \mathbb{Z}^3 \rightarrow \mathbb{R} \cup \{+\infty\}$  defined on  $\text{dom } f = [(0, 0, 0), (4, 2, 2)]_{\mathbb{Z}}$  by

$x_2$	$f(x_1, x_2, 0)$					$x_2$	$f(x_1, x_2, 1)$					$x_2$	$f(x_1, x_2, 2)$					
2	5	1	0	0	4	2	4	1	-2	-3	-1	2	6	3	0	-3	-4	
1	2	-1	-2	0	3	1	2	-1	-2	-3	-1	1	6	1	-2	-3	1	
0	0	-1	0	1	6	0	2	-1	-2	0	5	0	6	2	0	3	6	
	0	1	2	3	4	$x_1$	0	1	2	3	4	$x_1$	0	1	2	3	4	$x_1$

and let  $\alpha = 2$ . For  $x^\alpha = (0, 0, 0)$  we have  $f(x^\alpha) = 0$ ,  $f(x^\alpha) \leq f(x^\alpha + 2d)$  for  $d = (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 1, 1)$ . Hence  $x^\alpha = (0, 0, 0)$  is  $\alpha$ -local minimal. A unique (global) minimizer of  $f$  is located at  $x^* = (4, 2, 2)$  with  $f(x^*) = -4$  and  $\|x^\alpha - x^*\|_\infty = 4$ . The  $\ell_\infty$ -distance between  $x^\alpha$  and  $x^*$  is strictly larger than  $n(\alpha - 1) = 3$ . We remark that the scaled function  $f^\alpha$  is not integrally convex. ■

**Example 4.5.** For a positive integer  $m \geq 1$ , we consider two bipartite graphs  $G_1$  and  $G_2$  on vertex bipartition  $(\{0^+, 1^+, \dots, m^+\}, \{0^-, 1^-, \dots, m^-\})$ ; see Fig. 1. The edge sets of  $G_1$  and  $G_2$  are defined respectively as  $E_1 = \{(0^+, 0^-)\} \cup \{(i^+, j^-) \mid i, j = 1, \dots, m\}$  and  $E_2 = \{(0^+, j^-) \mid j = 1, \dots, m\} \cup \{(i^+, 0^-) \mid i = 1, \dots, m\}$ . Let  $V^+ = \{1^+, \dots, m^+\}$ ,  $V^- = \{1^-, \dots, m^-\}$ , and  $n = 2m + 2$ . Consider  $X_1, X_2 \subseteq \mathbb{Z}^n$  defined by

$$\begin{aligned} X_1 &= \left\{ \sum_{i=1}^m \sum_{j=1}^m \lambda_{ij}(\mathbf{1}_{i^+} - \mathbf{1}_{j^-}) + \lambda_0(\mathbf{1}_{0^+} - \mathbf{1}_{0^-}) \mid \begin{array}{l} \lambda_{ij} \in [0, \alpha - 1]_{\mathbb{Z}} \ (i, j = 1, \dots, m) \\ \lambda_0 \in [0, m^2(\alpha - 1)]_{\mathbb{Z}} \end{array} \right\}, \\ X_2 &= \left\{ \sum_{i=1}^m \mu_i(\mathbf{1}_{i^+} - \mathbf{1}_{0^-}) + \sum_{j=1}^m \nu_j(\mathbf{1}_{0^+} - \mathbf{1}_{j^-}) \mid \begin{array}{l} \mu_i \in [0, m(\alpha - 1)]_{\mathbb{Z}} \ (i = 1, \dots, m) \\ \nu_j \in [0, m(\alpha - 1)]_{\mathbb{Z}} \ (j = 1, \dots, m) \end{array} \right\}, \end{aligned}$$

where  $X_1$  and  $X_2$  represent the sets of boundaries of flows in  $G_1$  and  $G_2$ , respectively. We define functions  $f_1, f_2 : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  with  $\text{dom } f_1 = X_1$  and  $\text{dom } f_2 = X_2$  by

$$f_1(x) = \begin{cases} x(V^-) & (x \in X_1) \\ +\infty & (x \notin X_1), \end{cases} \quad f_2(x) = \begin{cases} x(V^-) & (x \in X_2) \\ +\infty & (x \notin X_2) \end{cases} \quad (x \in \mathbb{Z}^n),$$



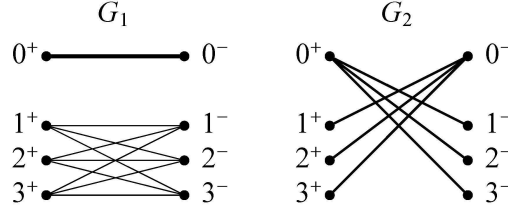


Figure 1: Example for  $O(n^2)$  lower bound for proximity distance ( $m = 3$ ).

where  $x(U) = \sum_{u \in U} x_u$  for any set  $U$  of vertices. Both  $f_1$  and  $f_2$  are M-convex, and hence  $f = f_1 + f_2$  is an  $M_2$ -convex function, which is integrally convex (see [23, Section 8.3.1]). We have  $\text{dom } f = \text{dom } f_1 \cap \text{dom } f_2 = X_1 \cap X_2$  and  $f$  is linear on  $\text{dom } f$ . As is easily verified,  $f$  has a unique minimizer at  $x^*$  defined by

$$x_u^* = \begin{cases} m(\alpha - 1) & (u \in V^+) \\ -m(\alpha - 1) & (u \in V^-) \\ m^2(\alpha - 1) & (u = 0^+) \\ -m^2(\alpha - 1) & (u = 0^-), \end{cases}$$

which corresponds to  $\lambda_0 = m^2(\alpha - 1)$ ,  $\lambda_{ij} = \alpha - 1$ ,  $\mu_i = \nu_j = m(\alpha - 1)$  ( $i, j = 1, \dots, m$ ). We mention that the function  $f$  here is constructed in [26, Remark 2.19] for a slightly different purpose (i.e., for  $M_2$ -proximity theorem).

Let  $x^\alpha = \mathbf{0}$ . Obviously,  $\mathbf{0} \in \text{dom } f$ . Moreover,  $x^\alpha = \mathbf{0}$  is  $\alpha$ -local minimal, since  $\text{dom } f \cap \{-\alpha, 0, \alpha\}^n = \{\mathbf{0}\}$ , as shown below. Since  $\|x^* - x^\alpha\|_\infty = m^2(\alpha - 1) = (n - 2)^2(\alpha - 1)/4$ , this example demonstrates a quadratic lower bound  $(n - 2)^2(\alpha - 1)/4$  for the proximity distance for integrally convex functions.

The proof of  $\text{dom } f \cap \{-\alpha, 0, \alpha\}^n = \{\mathbf{0}\}$  goes as follows. Let  $x \in X_1 \cap X_2 \cap \{-\alpha, 0, \alpha\}^n$ . We have  $x_{0^+} \in \{0, \alpha\}$  and  $x_{0^-} \in \{0, -\alpha\}$ . We consider four cases to conclude that  $x = \mathbf{0}$ .

- (i) Case of  $x_{0^+} = x_{0^-} = 0$ : The structure of  $X_2$  forces  $x = \mathbf{0}$ .
- (ii) Case of  $x_{0^+} = \alpha$ ,  $x_{0^-} = 0$ : The structure of  $X_2$  forces  $x_{i^+} = 0$  for  $i = 1, \dots, m$  and

$$x_{j^-} = \begin{cases} -\alpha & (j = j_0) \\ 0 & (j \neq j_0) \end{cases}$$

for some  $j_0$  ( $1 \leq j_0 \leq m$ ), but this is impossible by the structure of  $X_1$ .

- (iii) Case of  $x_{0^+} = 0$ ,  $x_{0^-} = -\alpha$ : We can similarly show that this is impossible.
- (iv) Case of  $x_{0^+} = \alpha$ ,  $x_{0^-} = -\alpha$ : The structure of  $X_2$  forces

$$x_{i^+} = \begin{cases} \alpha & (i = i_0) \\ 0 & (i \neq i_0), \end{cases} \quad x_{j^-} = \begin{cases} -\alpha & (j = j_0) \\ 0 & (j \neq j_0) \end{cases}$$

for some  $i_0$  ( $1 \leq i_0 \leq m$ ) and  $j_0$  ( $1 \leq j_0 \leq m$ ), but this is impossible by the structure of  $X_1$ . ■

We have seen that the proximity theorem with the linear bound  $n(\alpha - 1)$  does not hold for all integrally convex functions. Then a natural question arises: can we establish a proximity theorem at all by enlarging the proximity bound? This question is answered in the affirmative in Section 5.

## 4.2 A proximity theorem for integrally convex functions with $n = 2$

In the case of  $n = 2$  the proximity bound  $n(\alpha - 1) = 2(\alpha - 1)$  is valid for integrally convex functions<sup>2</sup>.

**Theorem 4.1.** *Let  $f : \mathbb{Z}^2 \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function,  $\alpha \in \mathbb{Z}_{++}$ , and  $x^\alpha \in \text{dom } f$ . If  $f(x^\alpha) \leq f(x^\alpha + \alpha d)$  for all  $d \in \{-1, 0, +1\}^2$ , then there exists a minimizer  $x^* \in \mathbb{Z}^2$  of  $f$  with  $\|x^\alpha - x^*\|_\infty \leq 2(\alpha - 1)$ .*

*Proof.* We may assume  $x^\alpha = \mathbf{0}$  and  $\alpha \geq 2$ . Define

$$C = \{x \in \mathbb{Z}^2 \mid 0 \leq x_2 \leq x_1\}, \quad S = \{x \in \mathbb{Z}^2 \mid 0 \leq x_2 \leq x_1 \leq 2(\alpha - 1)\}.$$

Let  $\mu$  be the minimum of  $f(x)$  over  $x \in S$  and let  $\hat{x}$  be a point in  $S$  with  $f(\hat{x}) = \mu$ . Then

$$f(x) \geq \mu \quad \text{for all } x \in S. \quad (4.1)$$

We will show that

$$f(2\alpha - 1, k) \geq \mu \quad (0 \leq k \leq 2\alpha - 1). \quad (4.2)$$

Then, by Corollary 2.5 (hyperplane-barrier property), it follows that  $f(z) \geq \mu$  for all  $z \in C$ , that is, there is no  $z \in C \setminus S$  with  $f(z) < \mu$ . This proves the claim of the theorem, since  $\mathbb{Z}^2$  can be covered by eight sectors similar to  $C$  and Proposition 2.1 holds.

The basic parallelogram inequality (3.4) with  $a = k$  and  $b = 2\alpha - 1 - k$  yields

$$f(0, 0) + f(2\alpha - 1, k) \geq f(k, k) + f(2\alpha - 1 - k, 0). \quad (4.3)$$

**Case 1:**  $0 \leq k \leq \alpha - 1$ . Since  $2\alpha - 1 - k \geq \alpha$ , by convexity of  $f(t, 0)$  in  $t$ , we have

$$\frac{1}{2\alpha - 1 - k} [f(2\alpha - 1 - k, 0) - f(0, 0)] \geq \frac{1}{\alpha} [f(\alpha, 0) - f(0, 0)] \geq 0.$$

On the other hand,  $f(k, k) \geq \mu$  by (4.1). Then it follows from (4.3) that

$$f(2\alpha - 1, k) \geq f(k, k) + [f(2\alpha - 1 - k, 0) - f(0, 0)] \geq \mu.$$

**Case 2:**  $\alpha \leq k \leq 2\alpha - 1$ . Since  $k \geq \alpha$ , by convexity of  $f(t, t)$  in  $t$ , we have

$$\frac{1}{k} [f(k, k) - f(0, 0)] \geq \frac{1}{\alpha} [f(\alpha, \alpha) - f(0, 0)] \geq 0.$$

On the other hand,  $f(2\alpha - 1 - k, 0) \geq \mu$  by (4.1). Then it follows from (4.3) that

$$f(2\alpha - 1, k) \geq f(2\alpha - 1 - k, 0) + [f(k, k) - f(0, 0)] \geq \mu.$$

We have thus shown (4.2), completing the proof of Theorem 4.1.  $\square$

## 5 A Proximity Theorem for Integrally Convex Functions

In this section we establish a proximity theorem for integrally convex functions in an arbitrary number of variables.

---

<sup>2</sup>Recall that  $n = 3$  in Example 4.4.

## 5.1 Theorem

**Theorem 5.1.** Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function,  $\alpha \in \mathbb{Z}_{++}$ , and  $x^\alpha \in \text{dom } f$ .

(1) If

$$f(x^\alpha) \leq f(x^\alpha + \alpha d) \quad (\forall d \in \{-1, 0, +1\}^n), \quad (5.1)$$

then  $\arg \min f \neq \emptyset$  and there exists  $x^* \in \arg \min f$  with

$$\|x^\alpha - x^*\|_\infty \leq \beta_n(\alpha - 1), \quad (5.2)$$

where  $\beta_n$  is defined by

$$\beta_1 = 1, \quad \beta_2 = 2; \quad \beta_n = \frac{n+1}{2}\beta_{n-1} + 1 \quad (n = 3, 4, \dots). \quad (5.3)$$

(2) The coefficient  $\beta_n$  of the proximity bound satisfies

$$\beta_n \leq \frac{(n+1)!}{2^{n-1}} \quad (n = 3, 4, \dots). \quad (5.4)$$

■

The numerical values of  $\beta_n$  and its bounds are as follows:

$n$	2	3	4	5	6	7
Value by (5.3)	2	5	13.5	41.5	146.25	586
Bound by (5.4)	–	6	15	45	157.5	630

(5.5)

**Remark 5.1.** The bound (5.2) can be strengthened to  $\|x^\alpha - x^*\|_\infty \leq \lfloor \beta_n(\alpha - 1) \rfloor$ , but  $\|x^\alpha - x^*\|_\infty \leq \lfloor \beta_n \rfloor(\alpha - 1)$  may not be correct (our proof does not justify this). ■

To prove Theorem 5.1 (1) we first note that the theorem follows from its special case where  $x^\alpha = \mathbf{0}$  and  $f$  is defined on a bounded set in the nonnegative orthant  $\mathbb{Z}_+^n$ . That is, the proof of Theorem 5.1 (1) is reduced to proving the following proposition. We use the notation  $N = \{1, 2, \dots, n\}$  and  $\mathbf{1}_A$  for the characteristic vector of  $A \subseteq N$ .

**Proposition 5.2.** Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function such that  $\text{dom } f$  is a bounded subset of  $\mathbb{Z}_+^n$  containing the origin  $\mathbf{0}$ . If

$$f(\mathbf{0}) \leq f(\alpha \mathbf{1}_A) \quad (\forall A \subseteq N), \quad (5.6)$$

then there exists  $x^* \in \arg \min f$  with

$$\|x^*\|_\infty \leq \beta_n(\alpha - 1), \quad (5.7)$$

where  $\beta_n$  is defined by (5.3). ■

Suppose that Proposition 5.2 has been established. Then Theorem 5.1 (1) can be derived from Proposition 5.2 in three steps:

1. We may assume  $x^\alpha = \mathbf{0}$  by Proposition 2.1 (1).

2. We may further assume that  $\text{dom } f$  is bounded. Let  $M$  be a sufficiently large integer, say,  $M \geq \beta_n(\alpha - 1) + 1$ , and  $f_M$  be the restriction of  $f$  to the integer interval  $[-M\mathbf{1}, M\mathbf{1}]_{\mathbb{Z}}$ , where  $\mathbf{1} = (1, 1, \dots, 1)$ . Then  $x^\alpha = \mathbf{0}$  is  $\alpha$ -local minimal for  $f_M$ . If the special case of Theorem 5.1 with  $x^\alpha = \mathbf{0}$  and bounded  $\text{dom } f$  is true, then there exists  $x^* \in \arg \min f_M$  satisfying  $\|x^*\|_\infty \leq \beta_n(\alpha - 1)$ . Since  $x^* \in \arg \min f_M$  we have  $f_M(x^*) \leq f_M(x^* + d)$  ( $\forall d \in \{-1, 0, +1\}^n$ ), which implies  $f(x^*) \leq f(x^* + d)$  ( $\forall d \in \{-1, 0, +1\}^n$ ). Then Theorem 2.3 shows that  $x^* \in \arg \min f$ .
3. We consider  $2^n$  orthants separately. For each  $s = (s_1, s_2, \dots, s_n) \in \{+1, -1\}^n$  we consider the function  $f_s(x) = f(sx)$  on  $\mathbb{Z}_+^n$ , where  $sx = (s_1x_1, s_2x_2, \dots, s_nx_n)$  and  $\text{dom } f_s = \mathbb{Z}_+^n$ , and apply Proposition 5.2 to  $f_s$  to obtain  $x_s^*$  with  $\|x_s^*\|_\infty \leq \beta_n(\alpha - 1)$ . From among  $2^n$  such  $x_s^*$ , take the one with the function value  $f(sx_s^*)$  minimum. Then  $x^* = sx_s^*$  is a minimizer of  $f$ , and satisfies  $\|x^*\|_\infty \leq \beta_n(\alpha - 1)$ .

## 5.2 Tools for the proof: $f$ -minimality

In this section we introduce some technical tools that we use in the proof of Proposition 5.2.

For  $A (\neq \emptyset) \subseteq N$ , we consider a set of integer vectors

$$B_A = \{\mathbf{1}_A + \mathbf{1}_i, \mathbf{1}_A - \mathbf{1}_i \mid i \in A\} \cup \{\mathbf{1}_A + \mathbf{1}_i \mid i \notin A\} \cup \{\mathbf{1}_A\}, \quad (5.8)$$

and the cones of their nonnegative integer and real combinations

$$C_A = \left\{ \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A \mid \mu_i^+, \mu_i^-, \mu_i^\circ, \lambda \in \mathbb{Z}_+ \right\}, \quad (5.9)$$

$$\tilde{C}_A = \left\{ \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A \mid \mu_i^+, \mu_i^-, \mu_i^\circ, \lambda \in \mathbb{R}_+ \right\}, \quad (5.10)$$

where  $C_A$  is often referred to as the integer cone generated by  $B_A$ . We first note the following fact, which provides us with a clearer geometric view, though it is not used in the proof of Proposition 5.2.

**Proposition 5.3.**  $B_A$  is a Hilbert basis of the convex cone  $\tilde{C}_A$  generated by  $B_A$ . That is,  $\tilde{C}_A = C_A \cap \mathbb{Z}^n$ .

*Proof.* The proof is given in Appendix B. □

For two nonnegative integer vectors  $x, y \in \mathbb{Z}_+^n$ , we write  $y \leq_f x$  if  $y \leq x$  and  $f(y) \leq f(x)$ . Note that  $y \leq_f x$  if and only if  $(y, f(y)) \leq (x, f(x))$  in  $\mathbb{R}^n \times (\mathbb{R} \cup \{+\infty\})$ . We say that  $x \in \mathbb{Z}_+^n$  is  $f$ -minimal if  $x \in \text{dom } f$  and there exists no  $y \in \mathbb{Z}_+^n$  such that  $y \leq_f x$  and  $y \neq x$ . That is<sup>3</sup>,  $x$  is  $f$ -minimal if and only if it is the unique minimizer of the function  $f$  restricted to the integer interval  $[\mathbf{0}, x]_{\mathbb{Z}}$ .

The goal of this section is to establish the following connection between  $f$ -minimality and the integer cone  $C_A$  based at  $\alpha\mathbf{1}_A$ .

**Proposition 5.4.** Assume  $\alpha$ -local minimality (5.6). If  $y \in \mathbb{Z}_+^n$  is  $f$ -minimal, then  $y \notin \alpha\mathbf{1}_A + C_A$  for any  $A (\neq \emptyset) \subseteq N$ . ■

Our proof of this proposition is based on several lemmas.

---

<sup>3</sup>  $x$  is  $f$ -minimal if and only if  $\arg \min f_{[\mathbf{0}, x]} = \{x\}$  for the function  $f_{[\mathbf{0}, x]}(y) = \begin{cases} f(y) & (y \in [\mathbf{0}, x]_{\mathbb{Z}}), \\ +\infty & (y \in \mathbb{Z}^n \setminus [\mathbf{0}, x]_{\mathbb{Z}}). \end{cases}$

**Lemma 5.5.** Assume  $\alpha$ -local minimality (5.6). For any  $A (\neq \emptyset) \subseteq N$  and  $\lambda \in \mathbb{Z}_+$  we have  $(\alpha - 1)\mathbf{1}_A \leq_f (\alpha - 1)\mathbf{1}_A + \lambda\mathbf{1}_A$ .

*Proof.* First note that  $(\alpha - 1)\mathbf{1}_A \leq (\alpha - 1)\mathbf{1}_A + \lambda\mathbf{1}_A$  for all  $\lambda \in \mathbb{Z}_+$ . By integral convexity of  $f$ ,  $g(\lambda) = f(\lambda\mathbf{1}_A)$  is a convex function in  $\lambda \in \mathbb{Z}_+$ , and therefore,

$$g(\alpha - 1) \leq \frac{\alpha - 1}{\alpha}g(\alpha) + \frac{1}{\alpha}g(0).$$

On the other hand,  $g(0) \leq g(\alpha)$  by the assumed  $\alpha$ -local minimality (5.6). Hence we have  $g(\alpha - 1) \leq g(\alpha)$ . Since  $g(0) < +\infty$ , by convexity of  $g$ , this implies  $g(\alpha - 1) \leq g((\alpha - 1) + \lambda)$  for all  $\lambda \in \mathbb{Z}_+$ , i.e.,  $f((\alpha - 1)\mathbf{1}_A) \leq f((\alpha - 1)\mathbf{1}_A + \lambda\mathbf{1}_A)$  for all  $\lambda \in \mathbb{Z}_+$ .  $\square$

**Lemma 5.6.** Let  $x \in \text{dom } f$ ,  $A (\neq \emptyset) \subseteq N$ , and assume  $x \leq_f x + \mathbf{1}_A$ . Then for any  $i \in N$ ,  $\delta \in \{+1, 0, -1\}$ , and  $\lambda \in \mathbb{Z}_+$  we have  $x + \mathbf{1}_A + \delta\mathbf{1}_i \leq_f (x + \mathbf{1}_A + \delta\mathbf{1}_i) + \lambda\mathbf{1}_A$ .

*Proof.* First note that  $x + \mathbf{1}_A + \delta\mathbf{1}_i \leq (x + \mathbf{1}_A + \delta\mathbf{1}_i) + \lambda\mathbf{1}_A$ . We only need to show  $f(x + \mathbf{1}_A + \delta\mathbf{1}_i) \leq f((x + \mathbf{1}_A + \delta\mathbf{1}_i) + \lambda\mathbf{1}_A)$  when  $f((x + \mathbf{1}_A + \delta\mathbf{1}_i) + \lambda\mathbf{1}_A) < +\infty$ . By integral convexity of  $f$  we have

$$\begin{aligned} & \frac{1}{\lambda + 1}f((x + \mathbf{1}_A + \delta\mathbf{1}_i) + \lambda\mathbf{1}_A) + \frac{\lambda}{\lambda + 1}f(x) \\ & \geq \tilde{f}(x + \mathbf{1}_A + \frac{\delta}{\lambda + 1}\mathbf{1}_i) \\ & = \frac{1}{\lambda + 1}f(x + \mathbf{1}_A + \delta\mathbf{1}_i) + \frac{\lambda}{\lambda + 1}f(x + \mathbf{1}_A), \end{aligned}$$

whereas  $f(x + \mathbf{1}_A) \geq f(x)$  by the assumption. Hence  $f((x + \mathbf{1}_A + \delta\mathbf{1}_i) + \lambda\mathbf{1}_A) \geq f(x + \mathbf{1}_A + \delta\mathbf{1}_i)$ .  $\square$

**Lemma 5.7.** Let  $x \in \text{dom } f$ ,  $A (\neq \emptyset) \subseteq N$ , and assume  $x \leq_f x + \mathbf{1}_A$ . For any  $\lambda \in \mathbb{Z}_+$ ,  $\mu_i^+, \mu_i^- \in \mathbb{Z}_+$  ( $i \in A$ ), and  $\mu_i^\circ \in \mathbb{Z}_+$  ( $i \notin A$ ), the point

$$y = x + \mathbf{1}_A + \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda\mathbf{1}_A \quad (5.11)$$

is not  $f$ -minimal.

*Proof.* By the definition of an  $f$ -minimal point, we assume  $y \in \text{dom } f$ ; since otherwise we are done. Define

$$\mu = \sum_{i \in A} (\mu_i^+ + \mu_i^-) + \sum_{i \notin A} \mu_i^\circ, \quad (5.12)$$

which serves, in our proof, as an index to measure the distance between  $x$  and  $y$ . If  $\mu \leq 1$ , then  $y$  is not  $f$ -minimal by Lemma 5.6. Suppose that  $\mu \geq 2$ . In the following we construct a vector  $x'$  such that  $x' \in \text{dom } f$ ,  $x' \leq_f x' + \mathbf{1}_A$ ,  $y$  is represented as (5.11) with  $x'$  in place of  $x$ , and the index  $\mu'$  for that representation is strictly smaller than  $\mu$ .

Define  $\beta = \mu + \lambda + 1$  and

$$\begin{aligned} A^+ &= \{i \in A \mid \mu_i^+ \geq 1\}, \\ A^- &= \{i \in A \mid \mu_i^- \geq 1\}, \\ A^\pm &= \{i \in A \mid \mu_i^+ = \mu_i^- = 0\}, \\ A^\circ &= \{i \in N \setminus A \mid \mu_i^\circ \geq 1\}, \end{aligned}$$

where we may assume, without loss of generality, that  $A^+ \cap A^- = \emptyset$ . Then (5.11) can be rewritten as

$$y = x + \sum_{i \in A^+} \mu_i^+ \mathbf{1}_i - \sum_{i \in A^-} \mu_i^- \mathbf{1}_i + \sum_{i \in A^\circ} \mu_i^\circ \mathbf{1}_i + \beta \mathbf{1}_A,$$

which shows

$$(y - x)_i = \begin{cases} \beta + \mu_i^+ & (i \in A^+), \\ \beta - \mu_i^- & (i \in A^-), \\ \beta & (i \in A^\circ), \\ \mu_i^\circ & (i \in A^\circ), \\ 0 & (\text{otherwise}). \end{cases}$$

Consider the point

$$z = \frac{\beta - 1}{\beta} x + \frac{1}{\beta} y.$$

Since  $f$  is integrally convex and  $x, y \in \text{dom } f$ , we have  $\tilde{f}(z) \leq ((\beta - 1)/\beta)f(x) + (1/\beta)f(y) < +\infty$ . On the other hand, since

$$(z - x)_i = \begin{cases} 1 + (\mu_i^+/\beta) & (i \in A^+), \\ 1 - (\mu_i^-/\beta) & (i \in A^-), \\ 1 & (i \in A^\circ), \\ \mu_i^\circ/\beta & (i \in A^\circ), \\ 0 & (\text{otherwise}), \end{cases}$$

the integral neighborhood  $N(z)$  of  $z$  consists of all points  $x'$  that can be represented as

$$x' = x + \mathbf{1}_A + \mathbf{1}_{A^+ \cap D} - \mathbf{1}_{A^- \cap D} + \mathbf{1}_{A^\circ \cap D} \quad (5.13)$$

for a subset  $D$  of  $A^+ \cup A^- \cup A^\circ$ . Since  $\tilde{f}(z) < +\infty$  and  $z \notin \mathbb{Z}^n$ , we must have  $|N(z) \cap \text{dom } f| \geq 2$ , which implies that there exists a nonempty  $D$  for which  $x' \in \text{dom } f$ . Take such  $D$  that is minimal with respect to set inclusion.

We claim that  $x' \preceq_f x' + \mathbf{1}_A$ . Obviously we have  $x' \leq x' + \mathbf{1}_A$ . To show  $f(x') \leq f(x' + \mathbf{1}_A)$ , we may assume  $x' + \mathbf{1}_A \in \text{dom } f$ . Then we have the following chain of inequalities:

$$\begin{aligned} & f(x) + f(x' + \mathbf{1}_A) \\ &= f(x) + f(x + 2\mathbf{1}_A + \mathbf{1}_{A^+ \cap D} - \mathbf{1}_{A^- \cap D} + \mathbf{1}_{A^\circ \cap D}) \\ &\geq 2\tilde{f}(x + \mathbf{1}_A + \frac{1}{2}\mathbf{1}_{A^+ \cap D} - \frac{1}{2}\mathbf{1}_{A^- \cap D} + \frac{1}{2}\mathbf{1}_{A^\circ \cap D}) && [\text{by integral convexity of } f] \\ &= f(x + \mathbf{1}_A) + f(x + \mathbf{1}_A + \mathbf{1}_{A^+ \cap D} - \mathbf{1}_{A^- \cap D} + \mathbf{1}_{A^\circ \cap D}) && [\text{by minimality of } D] \\ &= f(x + \mathbf{1}_A) + f(x') && [\text{by (5.13)}] \\ &\geq f(x) + f(x') && [\text{by } x \preceq_f x + \mathbf{1}_A] \end{aligned}$$

which shows  $f(x' + \mathbf{1}_A) \geq f(x')$ . Therefore,  $x' \preceq_f x' + \mathbf{1}_A$  is true.

We finally consider the index (5.12) associated with  $x'$ , which we denote by  $\mu'$ . The



substitution of (5.13) into (5.11) yields

$$\begin{aligned}
y &= x + \mathbf{1}_A + \sum_{i \in A^+} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^-} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \in A^\circ} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A \\
&= x' + \mathbf{1}_A + \left( \sum_{i \in A^+ \setminus D} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^+ \cap D} (\mu_i^+ - 1) (\mathbf{1}_A + \mathbf{1}_i) \right) \\
&\quad + \left( \sum_{i \in A^- \setminus D} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \in A^- \cap D} (\mu_i^- - 1) (\mathbf{1}_A - \mathbf{1}_i) \right) \\
&\quad + \left( \sum_{i \in A^\circ \setminus D} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^\circ \cap D} (\mu_i^\circ - 1) (\mathbf{1}_A + \mathbf{1}_i) \right) + (\lambda + |D| - 1) \mathbf{1}_A. \tag{5.14}
\end{aligned}$$

This shows  $\mu' = \mu - |D| \leq \mu - 1$ .

The above procedure finds  $x' \in \text{dom } f$  such that  $x' \leq_f x' + \mathbf{1}_A$  and  $\mu' \leq \mu - 1$ , when given  $x \in \text{dom } f$  such that  $x \leq_f x + \mathbf{1}_A$  and  $\mu \geq 2$ . By repeated application of this procedure we can eventually arrive at  $x'' \in \text{dom } f$  such that  $x'' \leq_f x'' + \mathbf{1}_A$  and  $\mu'' \leq 1$ . Then  $y$  is not  $f$ -minimal by Lemma 5.6 for  $x''$ .  $\square$

**Lemma 5.8.** *If  $(\alpha \mathbf{1}_A + C_A) \cap (\text{dom } f) \neq \emptyset$ , then  $\alpha \mathbf{1}_A \in \text{dom } f$ .*

*Proof.* To prove by contradiction, take a minimal  $y \in (\alpha \mathbf{1}_A + C_A) \cap (\text{dom } f)$  and assume that  $y \neq \alpha \mathbf{1}_A$ . The vector  $y$  can be represented as

$$y = \alpha \mathbf{1}_A + \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A,$$

where

$$\beta = \alpha + \sum_{i \in A^+} \mu_i^+ + \sum_{i \in A^-} \mu_i^- + \sum_{i \in A^\circ} \mu_i^\circ + \lambda$$

is strictly larger than  $\alpha$  since  $y \neq \alpha \mathbf{1}_A$ . Define

$$\begin{aligned}
A^+ &= \{i \in A \mid \mu_i^+ \geq 1\}, \\
A^- &= \{i \in A \mid \mu_i^- \geq 1\}, \\
A^\pm &= \{i \in A \mid \mu_i^+ = \mu_i^- = 0\}, \\
A^\circ &= \{i \in N \setminus A \mid \mu_i^\circ \geq 1\},
\end{aligned}$$

where we may assume, without loss of generality, that  $A^+ \cap A^- = \emptyset$ . We have

$$y_i = \begin{cases} \beta + \mu_i^+ & (i \in A^+), \\ \beta - \mu_i^- & (i \in A^-), \\ \beta & (i \in A^\pm), \\ \mu_i^\circ & (i \in A^\circ), \\ 0 & (\text{otherwise}). \end{cases}$$

Consider the point

$$z = \frac{\beta - 1}{\beta} y + \frac{1}{\beta} \mathbf{0}.$$

Since  $f$  is integrally convex and  $y, \mathbf{0} \in \text{dom } f$ , we have  $\tilde{f}(z) \leq ((\beta - 1)/\beta)f(y) + (1/\beta)f(\mathbf{0}) < +\infty$ . Note that

$$z_i = \begin{cases} (\beta - 1) + \mu_i^+ - \mu_i^+/\beta & (i \in A^+), \\ (\beta - 1) - \mu_i^- + \mu_i^-/\beta & (i \in A^-), \\ (\beta - 1) & (i \in A^\circ), \\ \mu_i^\circ - \mu_i^\circ/\beta & (i \in A^\circ), \\ 0 & (\text{otherwise}). \end{cases}$$

If  $A^+ \cup A^- \cup A^\circ = \emptyset$ , we are done with a contradiction. Indeed, we then have  $z = (\alpha + \lambda - 1)\mathbf{1}_A$  and  $y = (\alpha + \lambda)\mathbf{1}_A$ , and hence  $z \leq y$ ,  $z \neq y$ , and  $z \in \text{dom } f$  by  $f(z) = \tilde{f}(z) < +\infty$ . In the following we assume  $A^+ \cup A^- \cup A^\circ \neq \emptyset$ , which implies  $z \notin \mathbb{Z}^n$ .

The integral neighborhood  $N(z)$  of  $z$  consists of all points  $y'$  that can be represented as

$$y' = (\beta - 1)\mathbf{1}_A + \left( \sum_{i \in A^+ \setminus D} \mu_i^+ \mathbf{1}_i + \sum_{i \in A^+ \cap D} (\mu_i^+ - 1) \mathbf{1}_i \right) \\ - \left( \sum_{i \in A^- \setminus D} \mu_i^- \mathbf{1}_i + \sum_{i \in A^- \cap D} (\mu_i^- - 1) \mathbf{1}_i \right) + \left( \sum_{i \in A^\circ \setminus D} \mu_i^\circ \mathbf{1}_i + \sum_{i \in A^\circ \cap D} (\mu_i^\circ - 1) \mathbf{1}_i \right)$$

for a subset  $D$  of  $A^+ \cup A^- \cup A^\circ$ . Since  $\tilde{f}(z) < +\infty$  and  $z \notin \mathbb{Z}^n$ , we must have  $|N(z) \cap \text{dom } f| \geq 2$ , which implies that there exists a nonempty  $D$  for which  $y' \in \text{dom } f$ . Take any  $y' \in N(z) \cap \text{dom } f$  with  $D \neq \emptyset$ . Then  $y' \leq y$  and  $y' \neq y$ , since

$$y' - y = -\mathbf{1}_A - \sum_{i \in A^+ \cap D} \mathbf{1}_i + \sum_{i \in A^- \cap D} \mathbf{1}_i - \sum_{i \in A^\circ \cap D} \mathbf{1}_i \leq \mathbf{0}.$$

We also have  $y' \in (\alpha \mathbf{1}_A + C_A)$  by an alternative expression of  $y'$ :

$$y' = \alpha \mathbf{1}_A + \left( \sum_{i \in A^+ \setminus D} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^+ \cap D} (\mu_i^+ - 1) (\mathbf{1}_A + \mathbf{1}_i) \right) \\ + \left( \sum_{i \in A^- \setminus D} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \in A^- \cap D} (\mu_i^- - 1) (\mathbf{1}_A - \mathbf{1}_i) \right) \\ + \left( \sum_{i \in A^\circ \setminus D} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^\circ \cap D} (\mu_i^\circ - 1) (\mathbf{1}_A + \mathbf{1}_i) \right) + (\lambda + |D| - 1) \mathbf{1}_A.$$

Hence  $y' \in (\alpha \mathbf{1}_A + C_A) \cap (\text{dom } f)$ , a contradiction to the minimality of  $y$ .  $\square$

We are now in the position to prove Proposition 5.4. To prove the contrapositive of the claim, suppose that  $y \in \alpha \mathbf{1}_A + C_A$  for some  $A$ . Then  $y$  can be expressed as

$$y = \alpha \mathbf{1}_A + \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A$$

for some  $\mu_i^+, \mu_i^-, \mu_i^\circ, \lambda \in \mathbb{Z}_+$ . Equivalently,

$$y = ((\alpha - 1) \mathbf{1}_A + \mathbf{1}_A) + \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A,$$

which corresponds to the right-hand side of (5.11) with  $x = (\alpha - 1) \mathbf{1}_A$ . By Lemma 5.8, we have  $\alpha \mathbf{1}_A \in \text{dom } f$ . Since  $x = (\alpha - 1) \mathbf{1}_A \leq_f \alpha \mathbf{1}_A = x + \mathbf{1}_A$  by Lemma 5.5, Lemma 5.7 shows that  $y$  is not  $f$ -minimal. This completes the proof of Proposition 5.4.

### 5.3 Proof of Proposition 5.2 for $n = 2$

In this section we prove Proposition 5.2 for  $n = 2$  as an illustration of the proof method using the tools introduced in Section 5.2. This also gives an alternative proof of Theorem 4.1.

Recall that  $\text{dom } f$  is assumed to be a bounded subset of  $\mathbb{Z}_+^2$ , which implies, in particular, that  $\arg \min f \neq \emptyset$ . Take  $x^* = (x_1^*, x_2^*) \in \arg \min f$  that is  $f$ -minimal. We may assume  $x_1^* \geq x_2^*$  by Proposition 2.1 (2). Since  $x^*$  is  $f$ -minimal, Proposition 5.4 shows that  $x^*$  belongs to

$$X^* = \{(x_1, x_2) \in \mathbb{Z}_+^2 \mid x_1 \geq x_2\} \setminus ((\alpha \mathbf{1}_A + C_A) \cup (\alpha \mathbf{1}_N + C_N)),$$

where  $A = \{1\}$  and  $N = \{1, 2\}$ . On noting

$$\begin{aligned} C_A &= \{\mu_1(1, 0) + \mu_{12}(1, 1) \mid \mu_1, \mu_{12} \in \mathbb{Z}_+\}, \\ C_N &= \{\mu_1(1, 0) + \mu_2(0, 1) \mid \mu_1, \mu_2 \in \mathbb{Z}_+\}, \end{aligned}$$

we see that  $X^*$  consists of all integer points contained in the parallelogram with vertices  $(0, 0)$ ,  $(\alpha - 1, 0)$ ,  $(2\alpha - 2, \alpha - 1)$ ,  $(\alpha - 1, \alpha - 1)$ . Therefore,  $\|x^*\|_\infty \leq 2(\alpha - 1)$ . Thus Proposition 5.2 for  $n = 2$  is proved.

### 5.4 Proof of Proposition 5.2 for $n \geq 3$

In this section we prove Proposition 5.2 for  $n \geq 3$  by induction on  $n$ . Accordingly we assume that Proposition 5.2 is true for every integrally convex function in  $n - 1$  variables.

Let  $f : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  be an integrally convex function such that  $\text{dom } f$  is a bounded subset of  $\mathbb{Z}_+^n$  containing the origin  $\mathbf{0}$ . Note that  $\arg \min f \neq \emptyset$  and take  $x^* = (x_1^*, x_2^*, \dots, x_n^*) \in \arg \min f$  that is  $f$ -minimal. Then

$$[\mathbf{0}, x^*]_{\mathbb{Z}} \cap \arg \min f = \{x^*\}. \quad (5.15)$$

We may assume

$$x_1^* \geq x_2^* \geq \dots \geq x_n^* \quad (5.16)$$

by Proposition 2.1 (2).

The following lemma reveals a significant property of integrally convex functions that will be used here for induction on  $n$ . Note that, by (5.15),  $x^*$  satisfies the condition imposed on  $x^\bullet$ .

**Lemma 5.9.** *Let  $x^\bullet \in \text{dom } f$  be an  $f$ -minimal point. Then for any  $i \in N$  with  $x_i^\bullet \geq 1$  there exists an  $f$ -minimal point  $x^\circ \in \text{dom } f$  such that*

$$\mathbf{0} \leq x^\circ \leq x^\bullet, \quad \|x^\circ - x^\bullet\|_\infty = x_i^\bullet - x_i^\circ = 1.$$

*Proof.* Let  $x^\circ$  be a minimizer of  $f(x)$  among those  $x$  which satisfy the conditions:  $\mathbf{0} \leq x \leq x^\bullet$ ,  $\|x - x^\bullet\|_\infty = 1$ , and  $x_i = x_i^\bullet - 1$ ; in case of multiple minimizers, we choose a minimal minimizer with respect to the vector ordering (componentwise ordering). To prove  $f$ -minimality of  $x^\circ$ , suppose, to the contrary, that there exists  $z \in [\mathbf{0}, x^\circ]_{\mathbb{Z}} \setminus \{x^\circ\}$  with  $f(z) \leq f(x^\circ)$ . We have  $\ell = \|z - x^\bullet\|_\infty \geq 2$ .

Consider  $y \in \mathbb{R}_+^n$  defined by

$$y = \frac{\ell - 1}{\ell} x^\bullet + \frac{1}{\ell} z \quad (5.17)$$

and set  $Y = N(y) \cap \text{dom } f$ . The value of the local convex extension  $\tilde{f}$  of  $f$  at  $y$  can be represented as

$$\tilde{f}(y) = \sum_{y^j \in Y} \lambda_j f(y^j)$$

with some nonnegative coefficients  $\lambda_j$  such that

$$y = \sum_{y^j \in Y} \lambda_j y^j, \quad \sum_{y^j \in Y} \lambda_j = 1.$$

Since  $\|y - x^\bullet\|_\infty = 1$  and  $y \leq x^\bullet$  by (5.17), either  $y_i^j = x_i^\bullet - 1$  or  $y_i^j = x_i^\bullet$  holds for each  $y^j \in Y$ . Define

$$Y^< = \{y^j \in Y \mid y_i^j = x_i^\bullet - 1\}, \quad Y^= = \{y^j \in Y \mid y_i^j = x_i^\bullet\}.$$

Then we see

$$\sum_{y^j \in Y^<} \lambda_j = \frac{x_i^\bullet - z_i}{\ell}, \quad \sum_{y^j \in Y^=} \lambda_j = 1 - \frac{x_i^\bullet - z_i}{\ell} \quad (5.18)$$

from

$$y_i = \frac{\ell - 1}{\ell} x_i^\bullet + \frac{1}{\ell} z_i = x_i^\bullet - \frac{x_i^\bullet - z_i}{\ell}.$$

On the other hand, we have

$$\tilde{f}(y) \leq \frac{\ell - 1}{\ell} f(x^\bullet) + \frac{1}{\ell} f(z) \quad (5.19)$$

by integral convexity of  $f$ . We divide into cases to derive a contradiction to this inequality.

Case 1 ( $x_i^\bullet - z_i = \ell$ ): We have  $Y = Y^<$  and then  $f(x^\circ) \leq f(y^j)$  for all  $y^j \in Y$  by the definition of  $x^\circ$ . Hence

$$f(x^\circ) \leq \sum_{y^j \in Y} \lambda_j f(y^j) = \tilde{f}(y). \quad (5.20)$$

As for the right-hand side of (5.19), it follows from  $f(x^\bullet) < f(x^\circ)$  and  $f(z) \leq f(x^\circ)$  that

$$\frac{\ell - 1}{\ell} f(x^\bullet) + \frac{1}{\ell} f(z) < f(x^\circ). \quad (5.21)$$

But (5.20) and (5.21) together contradict (5.19).

Case 2 ( $x_i^\bullet - z_i < \ell$ ): In this case  $Y^=$  is nonempty. Since  $x^\bullet \notin Y$  by  $\|x^\bullet - y\|_\infty = 1$ , every  $y^j \in Y$  is distinct from  $x^\bullet$ , whereas  $y^j \in [0, x^\bullet]_{\mathbb{Z}}$ . Then the assumed  $f$ -minimality of  $x^\bullet$  implies

$$f(y^j) > f(x^\bullet) \quad (\forall y^j \in Y = Y^= \cup Y^<). \quad (5.22)$$

We also have

$$f(y^j) \geq f(x^\circ) \geq f(z) \quad (\forall y^j \in Y^<), \quad (5.23)$$

which is obvious from the definitions of  $x^\circ$  and  $z$ . Then we have

$$\begin{aligned}
\tilde{f}(y) &= \sum_{y^j \in Y^=} \lambda_j f(y^j) + \sum_{y^j \in Y^<} \lambda_j f(y^j) \\
&\quad \text{[by (5.22), (5.23), } Y^= \neq \emptyset \text{]} \\
&> \sum_{y^j \in Y^=} \lambda_j f(x^\bullet) + \sum_{y^j \in Y^<} \lambda_j f(z) \\
&\quad \text{[by (5.18)]} \\
&= (1 - \frac{x_i^\bullet - z_i}{\ell}) f(x^\bullet) + \frac{x_i^\bullet - z_i}{\ell} f(z) \\
&\quad \text{[by } \frac{x_i^\bullet - z_i}{\ell} \geq \frac{1}{\ell}, f(x^\bullet) \leq f(z) \text{]} \\
&\geq \frac{\ell - 1}{\ell} f(x^\bullet) + \frac{1}{\ell} f(z).
\end{aligned}$$

This is a contradiction to (5.19).  $\square$

Lemma 5.9 can be applied repeatedly, since the resulting point  $x^\circ$  satisfies the condition imposed on the initial point  $x^\bullet$ . Starting with  $x^\bullet = x^*$  we apply Lemma 5.9 repeatedly with  $i = n$ . After  $x_n^*$  applications, we arrive at a point  $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{n-1}, 0)$ . This point  $\hat{x}$  is  $f$ -minimal and

$$x_j^* - x_n^* \leq \hat{x}_j \quad (j = 1, 2, \dots, n-1). \quad (5.24)$$

We now consider a function  $\hat{f} : \mathbb{Z}^{n-1} \rightarrow \mathbb{R} \cup \{+\infty\}$  defined by

$$\hat{f}(x_1, x_2, \dots, x_{n-1}) = \begin{cases} f(x_1, x_2, \dots, x_{n-1}, 0) & (0 \leq x_j \leq \hat{x}_j \ (j = 1, 2, \dots, n-1)), \\ +\infty & (\text{otherwise}). \end{cases}$$

This function  $\hat{f}$  is an integrally convex function in  $n-1$  variables, and the origin  $\mathbf{0}$  is  $\alpha$ -local minimal for  $\hat{f}$ . By the induction hypothesis, we can apply Proposition 5.2 to  $\hat{f}$  to obtain

$$\|\hat{x}\|_\infty \leq \beta_{n-1}(\alpha - 1). \quad (5.25)$$

Note that  $\hat{x}$  is the unique minimizer of  $\hat{f}$ .

Combining (5.24) and (5.25) we obtain

$$x_1^* - x_n^* \leq \beta_{n-1}(\alpha - 1). \quad (5.26)$$

We also have

$$x_n^* \leq \frac{n-1}{n+1} x_1^* + \frac{2(\alpha-1)}{n+1} \quad (5.27)$$

as a consequence of  $f$ -minimality of  $x^*$ ; see Lemma 5.10 below. It follows from (5.26) and (5.27) that

$$x_1^* \leq x_n^* + \beta_{n-1}(\alpha - 1) \leq \frac{n-1}{n+1} x_1^* + \frac{2(\alpha-1)}{n+1} + \beta_{n-1}(\alpha - 1).$$

This implies

$$x_1^* \leq \left( \frac{n+1}{2} \beta_{n-1} + 1 \right) (\alpha - 1) = \beta_n(\alpha - 1),$$

where the recurrence relation

$$\beta_n = \frac{n+1}{2} \beta_{n-1} + 1$$

is used.

It remains to derive inequality (5.27) from  $f$ -minimality of  $x^*$ .

**Lemma 5.10.**

$$(1) \quad \sum_{i=1}^{n-1} (x_i^* - x_n^*) \geq x_n^* - \alpha + 1. \quad (5.28)$$

$$(2) \quad \sum_{i=2}^n (x_1^* - x_i^*) \geq x_1^* - \alpha + 1. \quad (5.29)$$

$$(3) \quad x_n^* \leq \frac{n-1}{n+1} x_1^* + \frac{2(\alpha-1)}{n+1}. \quad (5.30)$$

*Proof.* (1) To prove by contradiction, suppose that  $\sum_{i=1}^{n-1} (x_i^* - x_n^*) \leq x_n^* - \alpha$ . Then the expression

$$x^* = \alpha \mathbf{1}_N + \sum_{i=1}^{n-1} (x_i^* - x_n^*) (\mathbf{1}_N + \mathbf{1}_i) + \left( x_n^* - \alpha - \sum_{i=1}^{n-1} (x_i^* - x_n^*) \right) \mathbf{1}_N$$

shows  $x^* \in \alpha \mathbf{1}_N + C_N$ . By Proposition 5.4, this contradicts the fact that  $x^*$  is  $f$ -minimal.

(2) To prove by contradiction, suppose that  $\sum_{i=2}^n (x_1^* - x_i^*) \leq x_1^* - \alpha$ . Then the expression

$$x^* = \alpha \mathbf{1}_N + \sum_{i=2}^n (x_1^* - x_i^*) (\mathbf{1}_N - \mathbf{1}_i) + \left( x_1^* - \alpha - \sum_{i=2}^n (x_1^* - x_i^*) \right) \mathbf{1}_N$$

shows  $x^* \in \alpha \mathbf{1}_N + C_N$ . By Proposition 5.4, this contradicts the fact that  $x^*$  is  $f$ -minimal.

(3) Since

$$\sum_{i=1}^{n-1} (x_i^* - x_n^*) + \sum_{i=2}^n (x_1^* - x_i^*) = n(x_1^* - x_n^*),$$

the addition of (5.28) and (5.29) yields

$$n(x_1^* - x_n^*) \geq x_1^* + x_n^* - 2(\alpha - 1),$$

which is equivalent to (5.30). □

This completes the proof of Proposition 5.2, and hence that of Theorem 5.1 (1).

## 5.5 Estimation of $\beta_n$

The estimate of  $\beta_n$  given in Theorem 5.1 (2) is derived in this section.

The recurrence relation (5.3) can be rewritten as

$$\frac{2^n}{(n+1)!} \beta_n = \frac{2^{n-1}}{n!} \beta_{n-1} + \frac{2^n}{(n+1)!},$$

from which follows

$$\frac{2^n}{(n+1)!} \beta_n = \frac{2^2}{3!} \beta_2 + \sum_{k=3}^n \frac{2^k}{(k+1)!} = \frac{4}{3} + \sum_{k=3}^n \frac{2^k}{(k+1)!}. \quad (5.31)$$



For the last term we have

$$\sum_{k=3}^n \frac{2^k}{(k+1)!} \leq \sum_{k=3}^7 \frac{2^k}{(k+1)!} + \sum_{k=8}^{\infty} \frac{2^k}{(k+1)!} \leq \frac{167}{315} \quad (5.32)$$

since

$$\begin{aligned} \sum_{k=3}^7 \frac{2^k}{(k+1)!} &= \frac{2^3}{4!} + \frac{2^4}{5!} + \frac{2^5}{6!} + \frac{2^6}{7!} + \frac{2^7}{8!} = \frac{166}{315} (\approx 0.53), \\ \sum_{k=8}^{\infty} \frac{2^k}{(k+1)!} &\leq \frac{2^8}{9!} \sum_{k=1}^{\infty} \left(\frac{2}{10}\right)^{k-1} = \frac{320}{9!} (\approx 8.8 \times 10^{-4}) < \frac{1}{315}. \end{aligned}$$

Substitution of (5.32) into (5.31) yields

$$\beta_n \leq \left(\frac{4}{3} + \frac{167}{315}\right) \frac{(n+1)!}{2^n} = \frac{587}{315} \times \frac{(n+1)!}{2^n} \leq \frac{(n+1)!}{2^{n-1}}.$$

Thus the upper bound (5.4) is proved.

## 6 Optimization of Integrally Convex Functions

In spite of the facts that the factor  $\beta_n$  of the proximity bound is superexponential in  $n$  and that integral convexity is not stable under scaling, we can design a proximity-scaling type algorithm for minimizing integrally convex functions with bounded effective domains. The algorithm runs in  $C(n) \log_2 K_{\infty}$  time for some constant  $C(n)$  depending only on  $n$ , where  $K_{\infty}$  denotes the  $\ell_{\infty}$ -size of the effective domain. This means that, if the dimension  $n$  is fixed and treated as a constant, the algorithm is polynomial in the problem size. Note that no algorithm for integrally convex function minimization can be polynomial in  $n$ , since any function on the unit cube  $\{0, 1\}^n$  is integrally convex.

The proposed algorithm is a modification of the generic proximity-scaling algorithm given in the Introduction. In Step S1, we replace the function  $\tilde{f}(y) = f(x + \alpha y)$  with its restriction to the discrete rectangle  $\{y \in \mathbb{Z}^n \mid \|\alpha y\|_{\infty} \leq \beta_n(2\alpha - 1)\}$ , which is denoted by  $\hat{f}(y)$ . Then a local minimizer of  $\hat{f}(y)$  is found to update  $x$  to  $x + \alpha y$ . Note that a local minimizer of  $\hat{f}(y)$  can be found, e.g., by any descent method (the steepest descent method, in particular).

### Proximity-scaling algorithm for integrally convex functions

S0: Find an initial vector  $x$  with  $f(x) < +\infty$ , and set  $\alpha := 2^{\lceil \log_2 K_{\infty} \rceil}$ .

S1: Find an integer vector  $y$  that locally minimizes

$$\hat{f}(y) = \begin{cases} f(x + \alpha y) & (\|\alpha y\|_{\infty} \leq \beta_n(2\alpha - 1)), \\ +\infty & (\text{otherwise}), \end{cases}$$

in the sense of  $\hat{f}(y) \leq \hat{f}(y + d)$  ( $\forall d \in \{-1, 0, +1\}^n$ )

(e.g., by the steepest descent method), and set  $x := x + \alpha y$ .

S2: If  $\alpha = 1$ , then stop ( $x$  is a minimizer of  $f$ ).

S3: Set  $\alpha := \alpha/2$ , and go to S1.

### The steepest descent method to locally minimize $\hat{f}(y)$

D0: Set  $y := \mathbf{0}$ .

- D1: Find  $d \in \{-1, 0, +1\}^n$  that minimizes  $\hat{f}(y + d)$ .  
D2: If  $\hat{f}(y) \leq \hat{f}(y + d)$ , then stop ( $y$  is a local minimizer of  $\hat{f}$ ).  
D3: Set  $y := y + d$ , and go to D1.

The correctness of the algorithm can be shown as follows. We first assume that  $f$  has a unique (global) minimizer  $x^*$ . Let  $x^{2\alpha}$  denote the vector  $x$  at the beginning of Step S1, and define

$$f^\alpha(x) = \begin{cases} f(x) & (\|x - x^{2\alpha}\|_\infty \leq \beta_n(2\alpha - 1)), \\ +\infty & (\text{otherwise}), \end{cases} \quad \hat{f}^\alpha(y) = \begin{cases} f(x^{2\alpha} + \alpha y) & (\|\alpha y\|_\infty \leq \beta_n(2\alpha - 1)), \\ +\infty & (\text{otherwise}). \end{cases}$$

Note that  $f^\alpha$  is integrally convex, whereas  $\hat{f}^\alpha$  is not necessarily so. Let  $\hat{y}^\alpha$  be the output of Step S1 and  $x^\alpha = x^{2\alpha} + \alpha\hat{y}^\alpha$ . Then  $\hat{y}^\alpha$  is a local minimizer of  $\hat{f}^\alpha$  and  $x^\alpha - x^{2\alpha} = \alpha\hat{y}^\alpha \in (\alpha\mathbb{Z})^n$ .

**Lemma 6.1.**  $x^* \in \text{dom } f^\alpha$  for all  $\alpha$ .

*Proof.* This is obviously true in the initial phase with  $\alpha = 2^{\lceil \log_2 K_\infty \rceil}$ . To prove  $x^* \in \text{dom } f^\alpha$  by induction on descending  $\alpha$ , we show that  $x^* \in \text{dom } f^\alpha$  implies  $x^* \in \text{dom } f^{\alpha/2}$ . Since  $x^* \in \text{dom } f^\alpha$  and  $x^* \in \arg \min f$ , we have  $x^* \in \arg \min f^\alpha$ . On the other hand,  $x^\alpha$  is an  $\alpha$ -local minimizer of  $f^\alpha$ , since  $\hat{y}^\alpha$  is a local minimizer of  $\hat{f}^\alpha$ . Then, by the proximity theorem (Theorem 5.1) for  $f^\alpha$ , we obtain  $\|x^\alpha - x^*\|_\infty \leq \beta_n(\alpha - 1)$ , which shows  $x^* \in \text{dom } f^{\alpha/2}$ .  $\square$

In the final phase with  $\alpha = 1$ ,  $f^\alpha$  is an integrally convex function, and hence, by Theorem 2.3, an  $\alpha$ -local minimizer of  $f^\alpha$  is a global minimizer of  $f^\alpha$ . This observation, with Lemma 6.1 above, shows that the output of the algorithm is a global minimizer of  $f$ .

The complexity of the algorithm can be analyzed as follows. The number of iterations in the descent method is bounded by the total number of points in  $Y = \{y \in \mathbb{Z}^n \mid \|\alpha y\|_\infty \leq \beta_n(2\alpha - 1)\}$ , which is bounded by  $(4\beta_n)^n$ . For each  $y$  we examine all of its  $3^n$  neighboring points to find a descent direction or verify its local minimality. Thus Step S1, which updates  $x^{2\alpha}$  to  $x^\alpha$ , can be done with at most  $(12\beta_n)^n$  function evaluations. The number of scaling phases is  $\log_2 K_\infty$ . Therefore, the time complexity (or the number of function evaluations) is bounded by  $(12\beta_n)^n \log_2 K_\infty$ . For a fixed  $n$ , this gives a polynomial bound  $O(\log_2 K_\infty)$  in the problem size.

Finally, we describe how to get rid of the uniqueness assumption of the minimizer. Consider a perturbed function  $f_\varepsilon(x) = f(x) + \sum_{i=1}^n \varepsilon^i x_i$  with a sufficiently small  $\varepsilon > 0$ . By the assumed boundedness of the effective domain of  $f$ , the perturbed function has a minimizer, which is unique as a result of the perturbation. To find the minimum of  $f_\varepsilon$  it is not necessary to explicitly introduce parameter  $\varepsilon$  into the algorithm, but a lexicographically smallest local minimizer  $y$  of  $\hat{f}(y)$  should be found in Step S1.

**Remark 6.1.** Some technical points are explained here. By working with  $f^\alpha$ , we can bound the number of iterations for finding an  $\alpha$ -local minimizer in terms of the number of integer vectors contained in  $\text{dom } \hat{f}^\alpha$ . The vector  $x^\alpha$  is an  $\alpha$ -local minimizer for  $f^\alpha$ , but not necessarily for the original function  $f$ . This is why we apply the proximity theorem to  $f^\alpha$  in the proof of Lemma 6.1.  $\blacksquare$

**Remark 6.2.** The proximity bound  $\beta_n(\alpha - 1)$  in Theorem 5.1 is linear in  $\alpha$ . This linear dependence on  $\alpha$  is critical for the complexity  $O(\log_2 K_\infty)$  of the algorithm when  $n$  is fixed. Suppose, for example, that the proximity bound is  $\beta_n(\alpha^m - 1)$  for some  $m > 1$ . Then in the above analysis,  $(2\alpha - 1)$  should be replaced by  $((2\alpha)^m - 1)$ , and the total number of points

in  $Y = \{y \in \mathbb{Z}^n \mid \|ay\|_\infty \leq \beta_n((2\alpha)^m - 1)\}$  is bounded by  $(2^{m+1}\beta_n)^n \alpha^{(m-1)n}$ . The sum of  $\alpha^{(m-1)n}$  over  $\alpha = 1, 2, 2^2, \dots, 2^{\lceil \log_2 K_\infty \rceil}$  is of the order of  $K_\infty^{(m-1)n}$ . Then the proposed algorithm will not be polynomial in  $\log_2 K_\infty$ . Thus the particular form  $\beta_n(\alpha - 1)$  of our proximity bound is important for our algorithm. ■

## 7 Concluding Remarks

As shown in this paper, the nice properties of  $L^h$ -convex functions such as stability under scaling and the proximity bound  $n(\alpha - 1)$  are not shared by integrally convex functions in general. Two subclasses of integrally convex functions which still enjoy these nice properties have been introduced in [19] under the name of “directed integrally convex functions.” Directed integrally convex functions are defined as those functions which satisfy discrete midpoint convexity (1.1) for every pair  $(x, y) \in \mathbb{Z}^n \times \mathbb{Z}^n$  with  $\|x - y\|_\infty \geq 2$  or  $\|x - y\|_\infty = 2$ , where the former condition defines the “strong” version and the latter the “weak” version. Both classes of directed integrally convex functions are superclasses of  $L^h$ -convex functions, subclasses of integrally convex functions, are closed under scaling for all  $n$  and admit a proximity theorem with the bound  $n(\alpha - 1)$  for all  $n$ . See [19] for details.

## Acknowledgements

The authors thank Yoshio Okamoto for communicating a relevant reference. This research was initiated at the Trimester Program “Combinatorial Optimization” at Hausdorff Institute of Mathematics, 2015. This work was supported by The Mitsubishi Foundation, CREST, JST, and JSPS KAKENHI Grant Numbers 26350430, 26280004, 24300003, 16K00023.

## References

- [1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin: *Network Flows—Theory, Algorithms and Applications*, Prentice-Hall, 1993.
- [2] P. Favati and F. Tardella: Convexity in nonlinear integer programming, *Ricerca Operativa*, **53** (1990), 3–44.
- [3] S. Fujishige: Bisubmodular polyhedra, simplicial divisions, and discrete convexity, *Discrete Optimization*, **12** (2014), 115–120.
- [4] S. Fujishige and K. Murota: Notes on  $L$ -/ $M$ -convex functions and the separation theorems, *Mathematical Programming*, **88** (2000), 129–146.
- [5] R. Hemmecke, M. Köppe, J. Lee, and R. Weismantel: Nonlinear integer programming, in: M. Jünger et al., eds., *50 Years of Integer Programming 1958–2008*, Springer-Verlag, Berlin, 2010, Chapter 15, 561–618.
- [6] H. Hirai:  $L$ -extendable functions and a proximity scaling algorithm for minimum cost multiflow problem, *Discrete Optimization*, **18** (2015), 1–37.
- [7] H. Hirai:  $L$ -convexity on graph structures, arXiv:1610.02469 [math.OC], 2016.

- [8] D. S. Hochbaum: Complexity and algorithms for nonlinear optimization problems, *Annals of Operations Research*, **153** (2007), 257–296.
- [9] D. S. Hochbaum and J. G. Shanthikumar: Convex separable optimization is not much harder than linear optimization, *Journal of the Association for Computing Machinery*, **37** (1990), 843–862.
- [10] T. Ibaraki and N. Katoh: *Resource Allocation Problems: Algorithmic Approaches*, MIT Press, Boston, 1988.
- [11] T. Iimura: Discrete modeling of economic equilibrium problems, *Pacific Journal of Optimization*, **6** (2010), 57–64.
- [12] T. Iimura, K. Murota, and A. Tamura: Discrete fixed point theorem reconsidered, *Journal of Mathematical Economics*, **41** (2005), 1030–1036.
- [13] T. Iimura and T. Watanabe: Existence of a pure strategy equilibrium in finite symmetric games where payoff functions are integrally concave, *Discrete Applied Mathematics*, **166** (2014), 26–33.
- [14] S. Iwata, S. Moriguchi, and K. Murota: A capacity scaling algorithm for M-convex submodular flow, *Mathematical Programming*, **103** (2005), 181–202.
- [15] S. Iwata and M. Shigeno: Conjugate scaling algorithm for Fenchel-type duality in discrete convex optimization, *SIAM Journal on Optimization*, **13** (2003), 204–211.
- [16] N. Katoh, A. Shioura, and T. Ibaraki: Resource allocation problems, in: P.M. Pardalos, D.-Z. Du, R.L. Graham, eds., *Handbook of Combinatorial Optimization*, 2nd ed., Vol. 5, Springer-Verlag, Berlin, 2013, 2897–2988.
- [17] G. van der Laan, D. Talman, and Z. Yang: Solving discrete systems of nonlinear equations, *European Journal of Operational Research*, **214** (2011), 493–500.
- [18] S. Moriguchi, K. Murota, and A. Shioura: Scaling algorithms for M-convex function minimization. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, **E85-A** (2002), 922–929.
- [19] S. Moriguchi, K. Murota, A. Tamura, and F. Tardella: Integrally convex functions with discrete midpoint convexity, preprint, 2017.
- [20] S. Moriguchi, A. Shioura, and N. Tsuchimura: M-convex function minimization by continuous relaxation approach—Proximity theorem and algorithm, *SIAM Journal on Optimization*, **21** (2011), 633–668.
- [21] S. Moriguchi and N. Tsuchimura: Discrete L-convex function minimization based on continuous relaxation, *Pacific Journal of Optimization*, **5** (2009), 227–236.
- [22] K. Murota: Discrete convex analysis. *Mathematical Programming*, **83** (1998), 313–371.
- [23] K. Murota: *Discrete Convex Analysis*, SIAM, Philadelphia, 2003.

- [24] K. Murota: Recent developments in discrete convex analysis, in: W. Cook, L. Lovasz and J. Vygen, eds., *Research Trends in Combinatorial Optimization, Bonn 2008*, Springer-Verlag, Berlin, 2009, Chapter 11, 219–260.
- [25] K. Murota: Discrete convex analysis: A tool for economics and game theory, *Journal of Mechanism and Institution Design*, **1** (2016), 151–273.
- [26] K. Murota and A. Tamura: Proximity theorems of discrete convex functions, RIMS Preprint 1358, Kyoto University, June 2002.
- [27] K. Murota and A. Tamura: Proximity theorems of discrete convex functions, *Mathematical Programming*, **99** (2004), 539–562.
- [28] M. Queyranne and F. Tardella: Bimonotone linear inequalities and sublattices of  $\mathbb{R}^n$ , *Linear Algebra and Its Applications*, **413** (2006), 100–120.
- [29] A. Schrijver: *Theory of Linear and Integer Programming*, John Wiley and Sons, 1986.
- [30] A. Shioura: Fast scaling algorithms for M-convex function minimization with application to the resource allocation problem, *Discrete Applied Mathematics*, **134** (2003), 303–316.
- [31] A. Tamir: A strongly polynomial algorithm for minimum convex separable quadratic cost flow problems on series-parallel networks, *Mathematical Programming*, **59** (1993), 117–132.
- [32] A. Tamir: New pseudopolynomial complexity bounds for the bounded and other integer Knapsack related problems, *Operations Research Letters*, **37** (2009), 303–306.
- [33] A. Tamura: Coordinatewise domain scaling algorithm for M-convex function minimization, *Mathematical Programming*, **102** (2005), 339–354.
- [34] Z. Yang: On the solutions of discrete nonlinear complementarity and related problems, *Mathematics of Operations Research*, **33** (2008), 976–990.
- [35] Z. Yang: Discrete fixed point analysis and its applications, *Journal of Fixed Point Theory and Applications*, **6** (2009), 351–371.

## A An Alternative Proof of Theorem 2.2

Here is a proof of Theorem 2.2 (local characterization of integral convexity) that is shorter than the original proof in [2] and valid for functions defined on general integrally convex sets rather than discrete rectangles.

Obviously, (a) implies (b). The proof for the converse, (b)  $\Rightarrow$  (a), is given by the following two lemmas.

**Lemma A.1.** *Let  $B \subseteq \mathbb{R}^n$  be a box of size two with integer vertices, i.e.,  $B = [a, a + 2\mathbf{1}]_{\mathbb{R}}$  for some  $a \in \mathbb{Z}^n$ . If condition (b) is satisfied, then  $\tilde{f}$  is convex on  $B \cap \text{dom } f$ .*

*Proof.* We may assume  $B = [\mathbf{0}, 2\mathbf{1}]_{\mathbb{R}}$ . To prove by contradiction, assume that there exist  $x \in B \cap \text{dom } f$  and  $y^1, \dots, y^m \in B \cap \text{dom } f$  such that

$$x = \sum_{i=1}^m \lambda_i y^i, \quad \tilde{f}(x) > \sum_{i=1}^m \lambda_i f(y^i), \quad (\text{A.1})$$

where  $\sum_{i=1}^m \lambda_i = 1$  and  $\lambda_i > 0$  ( $i = 1, \dots, m$ ). We may also assume  $x \in [\mathbf{0}, \mathbf{1}]_{\mathbb{R}}$  without loss of generality. For each  $j = 1, \dots, n$ , we look at the  $j$ -th component of the generating points  $y^i$  to define

$$I_j^0 = \{i \mid y_j^i = 0\}, \quad I_j^2 = \{i \mid y_j^i = 2\}.$$

Since  $x_j = \sum_{i=1}^m \lambda_i y_j^i \leq 1$ , if  $I_j^2 \neq \emptyset$ , then  $I_j^0 \neq \emptyset$ .

Let  $j = n$  and suppose that  $I_n^2 \neq \emptyset$ . Then  $I_n^0 \neq \emptyset$ . We may assume  $y_n^1 = 0$ ,  $y_n^2 = 2$ ;  $\lambda_1 > 0$ ,  $\lambda_2 > 0$ . By (2.2) for  $(y^1, y^2)$  and the definition of  $\tilde{f}$  we have

$$f(y^1) + f(y^2) \geq 2\tilde{f}\left(\frac{y^1 + y^2}{2}\right) = 2 \sum_{k=1}^l \mu_k f(z^k),$$

where

$$\frac{y^1 + y^2}{2} = \sum_{k=1}^l \mu_k z^k, \quad z^k \in N\left(\frac{y^1 + y^2}{2}\right) \cap \text{dom } f \quad (k = 1, \dots, l) \quad (\text{A.2})$$

with  $\mu_k > 0$  ( $k = 1, \dots, l$ ) and  $\sum_{k=1}^l \mu_k = 1$ . This implies, with notation  $\lambda = \min(\lambda_1, \lambda_2)$ , that

$$\lambda_1 f(y^1) + \lambda_2 f(y^2) \geq (\lambda_1 - \lambda) f(y^1) + (\lambda_2 - \lambda) f(y^2) + 2\lambda \sum_{k=1}^l \mu_k f(z^k).$$

Hence

$$\sum_{i=1}^m \lambda_i f(y^i) \geq (\lambda_1 - \lambda) f(y^1) + (\lambda_2 - \lambda) f(y^2) + 2\lambda \sum_{k=1}^l \mu_k f(z^k) + \sum_{i=3}^m \lambda_i f(y^i).$$

Since

$$x = (\lambda_1 - \lambda) y^1 + (\lambda_2 - \lambda) y^2 + 2\lambda \sum_{k=1}^l \mu_k z^k + \sum_{i=3}^m \lambda_i y^i,$$

we have obtained another representation of the form (A.1). With reference to this new representation define  $\hat{I}_n^0$  (resp.,  $\hat{I}_n^2$ ) to be the set of indices of the generators whose  $n$ -th component is equal to 0 (resp., 2). Since  $z_n^k = 1$  for all  $k$  as a consequence of (A.2) with  $(y_n^1 + y_n^2)/2 = (0 + 2)/2 = 1$ , we have  $\hat{I}_n^0 \subseteq I_n^0$ ,  $\hat{I}_n^2 \subseteq I_n^2$  and  $|\hat{I}_n^0| + |\hat{I}_n^2| \leq |I_n^0| + |I_n^2| - 1$ .



By repeating the above process with  $j = n$ , we eventually arrive at a representation of the form of (A.1) with  $I_n^2 = \emptyset$ , which means that  $y_n^i \in \{0, 1\}$  for all generators  $y^i$ .

Then we repeat the above process for  $j = n - 1, n - 2, \dots, 1$ , to obtain a representation of the form of (A.1) with  $y^i \in [0, 1]_{\mathbb{Z}}$  for all generators  $y^i$ . This contradicts the definition of  $\tilde{f}$ .  $\square$

**Lemma A.2.** *For any  $x, y \in \overline{\text{dom } f}$ ,  $\tilde{f}$  is convex on the line segment connecting  $x$  and  $y$ .*

*Proof.* Let  $L$  denote the (closed) line segment connecting  $x$  and  $y$ , and consider boxes  $B$ , as in Lemma A.1, that intersect  $L$ . There exists a finite number of such boxes, say,  $B_1, \dots, B_m$ , and  $L$  is covered by the line segments  $L_j = L \cap B_j$  ( $j = 1, \dots, m$ ). That is,  $L = \bigcup_{j=1}^m L_j$ . For each point  $z \in L \setminus \{x, y\}$ , there exists some  $L_j$  that contains  $z$  in its interior, and  $\tilde{f}$  is convex on  $L_j$  by Lemma A.1. Hence<sup>4</sup>  $\tilde{f}$  is convex on  $L$ .  $\square$

## B Proof of Proposition 5.3

It is known (cf. [29, proof of Theorem 16.4]) that the set of integer vectors contained in

$$F_A = \left\{ \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A \left| \begin{array}{l} \mu_i^+, \mu_i^- \in [0, 1]_{\mathbb{R}} \ (i \in A); \\ \mu_i^\circ \in [0, 1]_{\mathbb{R}} \ (i \notin A); \\ \lambda \in [0, 1]_{\mathbb{R}} \end{array} \right. \right\}$$

forms a Hilbert basis of  $\tilde{C}_A$ . Let  $z$  be an integer vector in  $F_A$ . That is,  $z \in \mathbb{Z}^n$  and

$$z = \sum_{i \in A} \mu_i^+ (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A} \mu_i^- (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lambda \mathbf{1}_A \quad (\text{B.1})$$

$$= \sum_{i \in A} (\mu_i^+ - \mu_i^-) \mathbf{1}_i + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \left( \lambda + \sum_{i \in A} (\mu_i^+ + \mu_i^-) \right) \mathbf{1}_A \quad (\text{B.2})$$

for some  $\mu_i^+, \mu_i^- \in [0, 1]_{\mathbb{R}}$  ( $i \in A$ );  $\mu_i^\circ \in [0, 1]_{\mathbb{R}}$  ( $i \notin A$ );  $\lambda \in [0, 1]_{\mathbb{R}}$ . Our goal is to show that  $z$  can be represented as a nonnegative integer combination of vectors in  $B_A$ .

First note that  $\mu_i^\circ \in \{0, 1\}$  for each  $i \notin A$ ; define  $A^\circ = \{i \notin A \mid \mu_i^\circ = 1\}$ . We denote the coefficient of  $\mathbf{1}_A$  in (B.2) as

$$\xi = \lambda + \sum_{i \in A} (\mu_i^+ + \mu_i^-)$$

and divide into cases according to whether  $\xi$  is an integer or not.

Case 1 ( $\xi \in \mathbb{Z}$ ): Using  $\xi$  we rewrite (B.2) as

$$z = \sum_{i \in A} (\mu_i^+ - \mu_i^-) \mathbf{1}_i + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \xi \mathbf{1}_A,$$

in which  $\xi$  is an integer. For each  $i \in A$ ,  $\mu_i^+ - \mu_i^-$  must be an integer, which is equal to 0, 1 or  $-1$ . Accordingly we define

$$\begin{aligned} A^= &= \{i \in A \mid \mu_i^+ - \mu_i^- = 0\}, \\ A^> &= \{i \in A \mid \mu_i^+ - \mu_i^- = 1\} = \{i \in A \mid \mu_i^+ = 1, \mu_i^- = 0\}, \\ A^< &= \{i \in A \mid \mu_i^+ - \mu_i^- = -1\} = \{i \in A \mid \mu_i^+ = 0, \mu_i^- = 1\} \end{aligned}$$

<sup>4</sup>See H. Tuy: D.C. optimization: Theory, methods and algorithms, in: R. Horst and P. M. Pardalos, eds., *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht, 1995, 149–216; Lemma 2 to be specific.

to rewrite (B.1) as

$$z = \sum_{i \in A^>} (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^<} (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \in A^\circ} (\mathbf{1}_A + \mathbf{1}_i) + \left( \lambda + \sum_{i \in A^=} (\mu_i^+ + \mu_i^-) \right) \mathbf{1}_A. \quad (\text{B.3})$$

Here the coefficient of  $\mathbf{1}_A$  is integral, since

$$\lambda + \sum_{i \in A^=} (\mu_i^+ + \mu_i^-) = \xi - \sum_{i \in A^>} 1 - \sum_{i \in A^<} 1.$$

Hence (B.3) gives a representation of  $z$  as a nonnegative integer combination of vectors in  $B_A$ .

Case 2 ( $\xi \notin \mathbb{Z}$ ): Let  $\eta$  denote the fractional part of  $\xi$ , i.e.,  $\eta = \xi - \lfloor \xi \rfloor$  with  $0 < \eta < 1$ . We rewrite (B.2) as

$$z = \sum_{i \in A} (\mu_i^+ - \mu_i^- + \eta) \mathbf{1}_i + \sum_{i \notin A} \mu_i^\circ (\mathbf{1}_A + \mathbf{1}_i) + \lfloor \xi \rfloor \mathbf{1}_A. \quad (\text{B.4})$$

For each  $i \in A$ ,  $\mu_i^+ - \mu_i^- + \eta$  must be an integer, which is equal to 1 or 0. Accordingly we define

$$\begin{aligned} A^+ &= \{i \in A \mid \mu_i^+ - \mu_i^- + \eta = 1\}, \\ A^- &= \{i \in A \mid \mu_i^+ - \mu_i^- + \eta = 0\}. \end{aligned}$$

Then

$$\lfloor \xi \rfloor \geq \min(|A^+|, |A^-|),$$

which follows from

$$\begin{aligned} \mu_i^+ + \mu_i^- &\begin{cases} = 2\mu_i^- + 1 - \eta \geq 1 - \eta & (i \in A^+) \\ = 2\mu_i^+ + \eta \geq \eta & (i \in A^-), \end{cases} \\ \xi = \lambda + \sum_{i \in A} (\mu_i^+ + \mu_i^-) &\geq (1 - \eta)|A^+| + \eta|A^-| \geq \min(|A^+|, |A^-|). \end{aligned}$$

In the case of  $|A^+| \leq |A^-|$ , we see from (B.4) that

$$\begin{aligned} z &= \sum_{i \in A^+} \mathbf{1}_i + \sum_{i \in A^\circ} (\mathbf{1}_A + \mathbf{1}_i) + \lfloor \xi \rfloor \mathbf{1}_A \\ &= \sum_{i \in A^+} (\mathbf{1}_A + \mathbf{1}_i) + \sum_{i \in A^\circ} (\mathbf{1}_A + \mathbf{1}_i) + (\lfloor \xi \rfloor - |A^+|) \mathbf{1}_A, \end{aligned}$$

which is a nonnegative integer combination of vectors in  $B_A$ . In the other case with  $|A^+| > |A^-|$ , we have an alternative expression

$$\begin{aligned} z &= - \sum_{i \in A^-} \mathbf{1}_i + \sum_{i \in A^\circ} (\mathbf{1}_A + \mathbf{1}_i) + (\lfloor \xi \rfloor + 1) \mathbf{1}_A \\ &= \sum_{i \in A^-} (\mathbf{1}_A - \mathbf{1}_i) + \sum_{i \in A^\circ} (\mathbf{1}_A + \mathbf{1}_i) + (\lfloor \xi \rfloor + 1 - |A^-|) \mathbf{1}_A, \end{aligned}$$

which is also a nonnegative integer combination of vectors in  $B_A$ . This completes the proof of Proposition 5.3.